Summary

Chromatin and Epigenetic Regulation in Ground State Pluripotency

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät

der
Universität Zürich

von

Damian Dalcher

aus

Trubschachen BE

Promotionskommission

PD Dr. Raffaella Santoro (Vorsitz und Leitung der Dissertation)

Prof. Dr. Dr. Michael O. Hottiger

Prof. Dr. Constance Ciaudo

PD. Dr. Paolo Cinelli

Dr. Marcus Buschbeck

Zürich, 2018

Summary

Embryonic stem cells (ESCs) are characterized by their potential to self-renew indefinitely and by their ability to differentiate into all somatic cell types. This dual capacity makes ESCs a potential source for cell and tissue replacement therapies. However, before ESCs can safely and efficiently be used for clinical applications, intensive research on the mechanisms of pluripotency and ESC differentiation must be carried out. Chromatin and epigenetic regulatory mechanisms are well known to play key roles in ESC pluripotency. Remarkably, the chromatin of ESCs is particularly malleable and transcriptionally permissive. Such a property has been proposed to be crucial for the maintenance of a transcriptionally plastic state, which is required to enter in any lineage specification program during differentiation. ESCs exist in a variety of states in vitro, largely defined by the culture conditions, which reflect the natural development of the embryo from the blastocyst to post-implantation stages. When cultured in the presence of two small molecule inhibitors (ESCs+2i) closely resemble the developmental ground state in vivo whereas in the presence of serum and LIF (ESCs+serum), they are in a developmentally primed state. These two pluripotency states are defined by distinct gene expression and epigenetic signatures, which resemble cells of *in vivo* during pre- and post-implantation development, respectively. However, how chromatin structures and 3D genome organization is regulated during these different developmental states remains yet unclear. In this work, we aimed to identify the contribution of a particular chromatin remodeling complex called NoRC in maintaining ground state pluripotency. Although NoRC is highly expressed in both pluripotency states, only ESC+2i depend on NoRC for cell growth, differentiation, correct gene expression and H3K27me3 occupancy, whereas ESC+serum do not. Furthermore, we show that NoRC binds large and active regions within the A-compartment and limits far-cis chromatin contacts in order to maintain correct ground state 3D genome architecture. Strikingly, we identified two additional chromatin regulators, namely TOP2A and cohesin, which – in association with NoRC – also contribute to the integrity of ground state pluripotency. Particularly depletion of TOP2A or inhibition of TOP2 activities phenocopy the ground-state specific alterations observed upon TIP5 knockdown. These findings indicated that the ground-state chromatin structure is a fundamental player for correct gene expression and epigenetic states highlighting different degrees of genome dynamics during early phases of development.

Zusammenfassung

Embryonale Stammzellen (ESCs) zeichnen sich durch ihr Selbsterneuerungspotential und ihre Fähigkeit aus, sich in alle somatischen Zelltypen zu differenzieren. Diese doppelte Kapazität macht ESC zu einer potenziellen Quelle für Zell- und Gewebeersatztherapien. Bevor jedoch ESCs sicher und effizient für klinische Anwendungen eingesetzt werden können, muss intensiv nach den Mechanismen der Pluripotenz und der ESC-Differenzierung geforscht werden. Chromatin und epigenetische Regulationsmechanismen spielen bekanntermaßen eine Schlüsselrolle bei der Pluripotenz der ESCs. Bemerkenswerterweise ist das Chromatin von ESCs besonders formbar und transkriptionell permissiv. Es wurde vorgeschlagen, dass eine solche Eigenschaft für die Aufrechterhaltung eines transkriptionell plastischen Zustands entscheidend ist, welcher für die Differenzierung in jedes Linienspezifikationsprogramm gebraucht wird. ESCs existieren in einer Vielzahl von Zuständen *in vitro*, die weitgehend durch die Kulturbedingungen definiert sind und die natürliche Entwicklung des Embryos von der Blastozyste bis zur Post-Implantationsphase widerspiegeln. Werden sie in Gegenwart von zwei kleinen Molekülinhibitoren (ESCs+2i) gezüchtet, ähneln sie dem Entwicklungsgrundzustand *in vivo*, während sie in Anwesenheit von Serum und LIF (ESCs+serum) in einem entwicklungsgeprimten Zustand gehalten werden. Diese zwei Pluripotenzzustände sind durch unterschiedliche Genexpression und epigenetische Signaturen definiert, die Zellen der *in vivo* Prä- und Post-Implantationsentwicklung ähneln. Wie Chromatinstrukturen und die 3D-Genomorganisation während den verschiedenen Entwicklungsstadien reguliert werden, ist jedoch noch unklar. In dieser Arbeit wollten wir den Beitrag eines bestimmten Chromatin-Remodelling-Komplexes namens NoRC zur Aufrechterhaltung der Pluripotenz im Grundzustand analysieren. Obwohl NoRC in beiden Pluripotenzzuständen stark exprimiert wird, hängen nur ESCs+2i für Zellwachstum, Differenzierung, korrekte Genexpression und H3K27me3-Besetzung von NoRC ab, während ESCs+serum dies nicht tun. Darüber hinaus zeigen wir, dass NoRC grosse und aktive Regionen innerhalb des A-Kompartiments bindet und weite cis Chromatinkontakte begrenzt, um korrekte 3D-Genomarchitektur des Grundzustands zu erhalten. Bemerkenswerterweise identifizierten wir zwei zusätzliche Chromatinregulatoren, nämlich TOP2A und Cohesin, die zusammen mit NoRC auch zur Integrität der Pluripotenz im Grundzustand beitragen. Insbesondere

der Verlust von TOP2A oder seine Hemmung durch spezifische Inhibitoren zeigen die nach TIP5-Knockdown beobachteten Grundzustands-spezifischen Veränderungen. Diese Befunde zeigten, dass die Chromatinstruktur im Grundzustand ein fundamentaler Faktor für korrekte Genexpression und epigenetische Zustände ist, die unterschiedliche Grade der Genomdynamik in frühen Phasen der Embryonalentwicklung hervorheben.

Abbreviations

2i	2 inhibitor
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
ac	acetylation
ATP	adenosine triphosphate
bp	base pair
CGI	CpG island
CHD	chromodomain helicase DNA-binding
ChIP	chromatin immunoprecipitation
CT	chromosome territory
CTCF	CCCTC-binding factor
DamID	DNA adenine methyltransferase identification
DNA	desoxyribonucleic acid
DNMT	DNA methyltransferase
EpiSCs	epiblast stem cell
ERK	extracellular signal-regulated kinase
ESC	embryonic stem cell
GSK3	glycogen synthase kinase 3
HAT	histone acetyl transferase
HDAC	histone de-acetylase
HiC	chromatin conformation capture sequencing
HMT	histone methyl transferase
ICM	inner cell mass
IGS-rRNA	intergenic spacer ribosomal RNA
iPSCs	induced pluripotent stem cells
ISWI	imitation switch
SWI/SNF	switch/ sucrose non-fermentable
JAK	Janus kinase
kb	kilo bases
KO	knock out
LAD	lamina-associated domain
LIF	leukemia inhibitory factor

lncRNA	long non coding RNA
Mb	mega bases
me	methylation
MEK	mitogen-activated protein kinase 1
NAD	nucleolus-associated domain
NL	nuclear lamina
NOR	nucleolus organizing region
NoRC	nuclelar remodeling complex
nt	nucleotides
OCT_4	octamer-binding transcription factor 4
OSN	OCT4/SOX2/NANOG
PGC	primordial germ cell
ph	phosphorylation
Pramel7	PRAME-like 7
PRC	Polycomb repressive complex
pRNA	promoter RNA
PTM	post-translational modification
rDNA	ribosomal DNA
Rex1	Reduced Expression Protein 1
RNAi	RNA interference
RNP	ribonucleoprotein
rRNA	ribosomal RNA
siRNA	small interfering RNA
SMC	structural maintenance of chromosomes
SNF2H	sucrose nonfermenting protein 2 homolog
SOX2	SRY-Box 2
STAT3	signal transducer and activator of transcription 3
T	Terminator element
TAD	topological-associated domain
TDG	thymine DNA glycosylase
TET	ten-eleven translocation
TIP5	TTF1 interacting protein 5
TKO	triple knock out

TOP2A	topoisomerase 2a
tRNA	transfer RNA
TSS	transcription start site
TTF1	transcription termination factor 1
ub	ubiquitination
UBF	upstream binding factor
UCE	upstream control element
UHRF1	ubiquitin like with PHD and ring finger domains 1
ZGA	zygotic genome activation

Table of Contents

1. Introduction

1.1 Embryonic stem cells (ESCs)

Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of a blastocyst, an early pre-implantation stage of the developing embryo (in mouse, E3.5-E4.5). Under optimal culturing conditions, these cells can be maintained in a pluripotent state that is characterized by the potential to self-renew indefinitely and by the ability to differentiate into all three germ layers endoderm, mesoderm and ectoderm (Martello and Smith, 2014). ESCs are of great interest as a model system for studying early developmental processes and for their potential therapeutic application in regenerative medicine (Weinberger et al., 2016). Considering the similarities of ESCs and cancer cells, extending the basic knowledge in stem cells might also lead to new therapeutic approaches for cancer therapies (Gaspar-Maia et al., 2011a).

1.1.1 Origin of mouse ESCs

In 1981 Evans and Martin independently succeeded for the first time in deriving pluripotent ESCs directly from the blastocyst of the developing mouse embryo. Co-culturing these cells in a serum-containing medium with a feeder layer of mitotically inactivated fibroblasts allowed the maintenance of an undifferentiated state (Evans and Kaufman, 1981; Martin, 1981). Their pluripotent capacity was confirmed by two independent experimental readouts. Firstly, they generated teratocarcinomas that are composed of multiple tissue types from all three germ layers when grafted to adult mice. Secondly, these cells were capable of contributing to the germ line of chimeric animals which themselves were fertile and generated healthy offspring (Bradley et al., 1984). Since ESCs were permissive to standard genetic manipulations and could be clonally expanded, ESCs could be used as a vehicle for introducing mutations into mice. In 1989, the first mouse knock out (KO) mouse was generated by gene targeting in ESCs (Thompson et al., 1989). Although *in vitro* culturing of ESCs was doubtlessly a break through in the field of stem cell research, the efficiency of ESC derivation and cultivation was very low. A large improvement was achieved by the addition of leukeamia inhibitory factor (LIF) to the culturing media. LIF was identified as a secreted factor from feeder cells that could effectively suppress

differentiation of ESCs (Smith et al., 1988; Williams et al., 1988). LIF signals to activate the JAK-STAT3 (Janus kinase-signal transducer and activator of transcription 3) pathway, which controls the transcription of genes regulating self-renewal such as KLF4, GBX2 and C-MYC (**Figure 1**) (Martello and Smith, 2014). Thus, LIF could be used to replace feeders for both derivation and propagation of ESCs. However, LIF alone was not sufficient for self-renewal, since removal of serum from the medium led to a gradual loss in ESC pluripotency. This obstacle was overcome in 2003 by the discovery of bone morphogenic proteins (BMPs) as potent antagonists of neural specification in vertebrate embryos. Strikingly, BMP4 was found to be able to replace serum and – in combination with LIF – sustain ESC self-renewal in a serum- and feeder-free condition (Ying et al., 2003). Mechanistically, BMP4 signals via SMAD proteins to activate Inhibitor of Differentiation (ID) genes (**Figure 1**) (Malaguti et al., 2013).

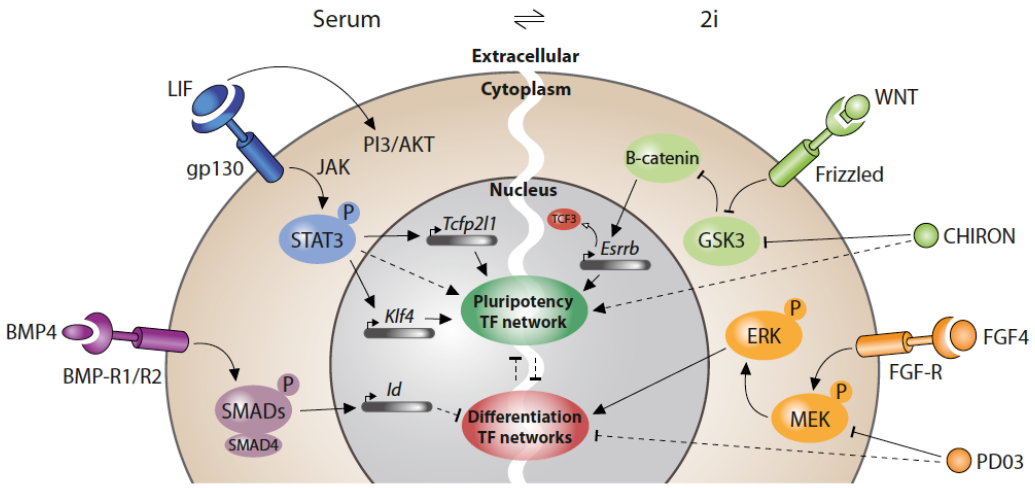


Figure 1. Extrinsic Signaling Pathways reinforcing or antagonizing self-renewal. ESCs+serum depend on BMP4 and LIF signaling via SMAD and STAT3, respectively. In ESCs+2i ChIRON inhibits GSK3, which mimics canonical WNT signaling and results in activation of Esrrb expression. Inhibition of ERK signaling via PD interferes with FGF4 signaling to activate differentiation related genes. Taken from **(Hackett and Surani, 2014)**

1.1.2 Ground state, naïve and primed pluripotency

The ability of maintaining ESCs in culture resulted in the definition of naïve pluripotency, which describes the property of cells giving rise to all embryonic lineages following blastocyst injection (Nichols and Smith, 2009). Although conventional ESCs cultured in serum+LIF (ESCs+serum) meet the rather descriptive criteria of naïve pluripotency, they have been shown to be heterogeneous and differ

substantially from the *in vivo* epiblast cells in the ICM. Particularly, the molecularly undefined serum component is known to contribute to a conflict in the pluripotency network. The interplay of pro- and anti-differentiation signaling pathways induce a metastable state leading to significant transcriptional, morphological and functional heterogeneity among these cells (Chambers et al., 2007; Hayashi et al., 2008; Toyooka et al., 2008). As a result, ESCs+serum exist in at least two distinct populations. One corresponding to the naïve pluripotent cells, while the other subpopulation resembles a developmentally more advanced state associated with enhanced expression of lineage-specific genes (**Figure 2**). The later population has further been shown to perform poorly in pluripotency assays (Hackett and Surani, 2014). Interestingly, ESCs+serum were shown to be interchangeable between both subpopulations resulting in a dynamic and metastable state allowing cycling in and out of naïve pluripotency (Abranches et al., 2013). The problematic heterogeneity of ESCs+serum was overcome in 2008 by supplementing the ESC medium with two small molecule inhibitors referred to as "2i". Blocking GSK3 and MEK/ERK activities by CHIR99021 and PD0325901, respectively resulted in the trapping of the ESCs in their naïve pluripotent state (ESCs+2i) (**Figure 1**) (Ying et al., 2008). Both GSK3 and MEK/ERK signaling are known to be antagonistic with self-renewal (Doble et al., 2007; Kunath et al., 2007; Sato et al., 2004). Mechanistically, EKR2, that is activated upon FGF4 signaling, phosphorylates the pluripotency protein KLF4 leading to its proteasomal degradation (Yeo et al., 2014). Consequently, the pluripotency network is stabilized by the inhibition of MEK, the upstream activator of ERK1/2 (**Figure 1**) (Ying et al., 2008). The beneficial impact of GSK3 inhibition on self-renewal is mainly mediated by stabilization of β-catenin and therefore mimicking stimulation of the canonical WNT signaling pathway (**Figure 1**) (ten Berge et al., 2011; Yi et al., 2011). In the absence of β-catenin the transcriptional repressor TCF3 associates pluripotency genes and antagonizes with OCT4 and SOX2 activity. Upon stabilization, β-catenin translocates to the nucleus and interacts with TCF3 resulting in the disruption of TCF3-mediated repression (Faunes et al., 2013; Wray et al., 2011). ESRRB has been shown to be among the most important factors upregulated in response to CHIR99021 treatment and thereby release from TCF3-mediated repression (**Figure 1**) (Martello et al., 2012). In summary, addition of these two inhibitors allows the derivation and propagation of ESCs in the absence of serum, which can therefore be maintained in a chemically well-defined medium allowing

more reproducible results. 2i-culturing conditions increases the ESC derivation efficiency and also allows ESC retrieval of a broader species spectrum. While ESCs+serum are successfully isolated from 129 and hybrid mouse strains only, ESCs+2i are derived from all mouse strains and even from other rodent species, such as from rat (Buehr et al., 2008; Kiyonari et al., 2010; Li et al., 2008; Nichols et al., 2009). Molecular characterization of ESCs+2i revealed a highly homogenous population of cells that seem to be isolated from perturbation by differentiation inducers. Moreover, ESCs+2i exhibit reduced expression of lineage-related genes, a highly permissive epigenetic landscape and cluster closely with E4.5 epiblast cells at the single-cell transcriptome level. By contrast, ESCs+serum appear to resemble developmentally more advanced state of differentiation, as assessed by analyses of their transcriptome and epigenome (Boroviak et al., 2014; Marks et al., 2012). Although both ESCs+serum and ESCs+2i are functionally naïve pluripotent, ESCs+2i were additionally defined to be in a ground state that is defined as unrestricted naïve pluripotency established *in vivo* in the epiblast cells of the developing blastocyst (**Figure 2**). Ground state pluripotency of ESCs+2i is further supported by the fact that cells, that have left the naïve pluripotent state, usually die in the presence of 2i (Silva and Smith, 2008). Taken together, ESCs+2i are referred to as ground state pluripotent and are to date the best *in vitro* approximation of the epiblast cells from the developing embryo.

Besides from the ICM, pluripotent stem cells can also be isolated from the post-implantation epiblast cells at E5.5-8.0. The so-called epiblast stem cells (EpiSCs) are able to form tertacarcinomas and are capable of differentiating into all germ layers *in vitro*. However, upon injection into recipient mice they do not or only very inefficiently contribute to chimeric animals (Brons et al., 2007; Hayashi et al., 2011). This restricted or "primed" pluripotency is one of the main functional characteristics of EpiSCs representing an even more advanced developmental stage. Different than naïve ESCs, EpiSC require FGF2 and Activin A signaling for maintaining self-renewal (**Figure 2**) (Brons et al., 2007). Transcriptionally, EpiSCs show reduced expression of many ancillary pluripotency factors (Festuccia et al., 2012). Furthermore, EpiSCs accumulate epigenetic barriers such as female X chromosome inactivation and promoter methylation of pluripotency genes. These molecular characteristics seem to be incompatible with the naïve state. Accordingly, EpiSCs are

distinct from ESCs+serum/ESCs+2i and functionally, they rather resemble the anterior primitive streak (Kojima et al., 2014).

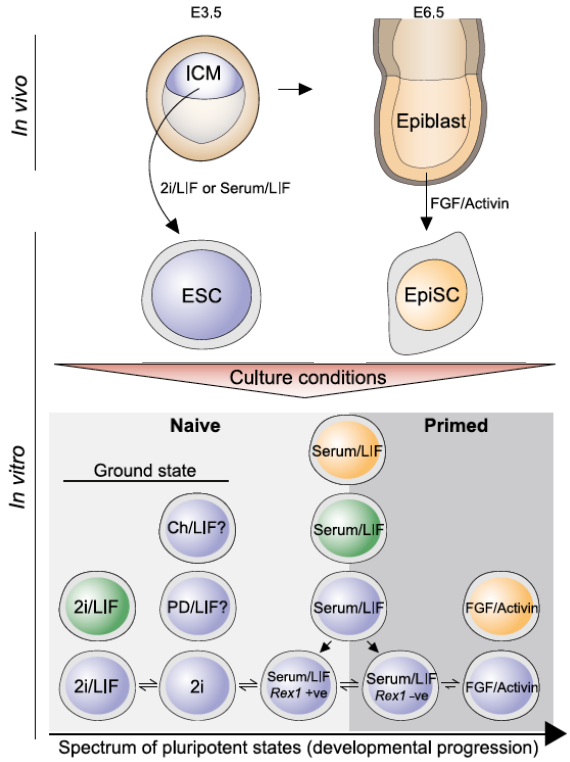


Figure 2. Spectrum of ground state and primed pluripotent stem cells. ESCs are isolated from the blastocyst, while EpiSCs derive from post-implantation epiblast. Depending on the derivation and culturing condition, these pluripotent stem cells reside in distinct pluripotency states, which are classified as naïve and primed. Furthermore, an unrestricted naïve state closely recapitulating the epiblast cells in the ICM is additionally defined as ground state. ESCs+2i and EpiSCs are classified as ground state and primed, respectively. ESCs+serum however, exist in a highly heterogeneous population of naïve and primed ESCs, which is reflected by high and low Rex1 expression. Taken and modified from (Hackett and Surani, 2014)

Taken together, three main pluripotent states are distinguished among the cultured embryonic stem cells *in vitro*: (i) ground state (ii) naïve and (iii) primed pluripotency that are dependent on the ESC culturing conditions and on the developmental state of the embryo upon retrieval. Although these states are molecularly and functionally distinct, it is possible to interchange from to another and therefore model *in vitro* processes that occur during early embryonic development. An exact definition of ESCs+serum has been challenging due to their heterogeneity within the population. While one part is functionally naïve and highly expresses pluripotency factors, the other subpopulation performs poorly in pluripotency assays and seemed to be in a more primed state (Hackett and Surani, 2014). Molecular profiling of ESCs+serum revealed that these cells – in contrast to ESCs+2i – transcriptionally and

epigenetically resemble a more advanced post-implantation state of development. Even the most naïve ESCs+serum, as judged by high expression of Rex1, appear developmentally more advanced than ground state ESCs+2i and therefore represent a distinct molecular state of naïve pluripotency (**Figure 2**) (Marks et al., 2012). For the future of this work ESCs+serum will therefore be referred to as "primed", while ESCs+2i have doubtlessly been shown to reside in ground state pluripotency.

1.1.3 The pluripotency gene regulatory network

Downstream of the extrinsic signals, a complex network of transcription factors activates self-renewal related genes and represses differentiation-associated pathways. This intrinsic genetic network is organized into three "core" and many "ancillary" pluripotency factors. Octamer-binding transcription factor 4 (OCT4 or POU5F1) is uniquely expressed in all ESCs and is absolutely essential for *in vivo* and *in vitro* pluripotency. It belongs to the Pic-Oct_Unc (POU) family of transcription factors and is exclusively expressed in ESCs and primordial germ cells (Rosner et al., 1990; Scholer et al., 1990). Similarly, SRY-box 2 (SOX2) is indispensable for the formation of the pluripotent epiblast. It is a member of the high mobility group (HMG) DNA binding proteins and among others a key regulator of OCT4 gene expression (Avilion et al., 2003). In contrast to OCT4, SOX2 is more broadly expressed in mouse. In addition to the pre- and postimplantation epiblast cells, SOX2 is present in trophectoderm and in the all neuroectodermal lineages. Interestingly, both depletion and overexpression OCT4 and SOX2 lead to ESC differentiation, highlighting that tight control of OCT4 and SOX2 expression is the foundation for the pluripotency gene regulatory network (Niwa et al., 2000). NANOG, a homeodomain-containing transcription factor, is considered as the third core pluripotency factor due to its crucial role in establishing the pluripotent state in the ICM and because of its ability to maintain pluripotency in the absence of LIF when overexpressed in ESCs (Chambers et al., 2007; Mitsui et al., 2003). However, ESCs are still capable to self-renew upon its genetic depletion, implying that NANOG executes a different functional role from OCT4 and SOX2 (Carter et al., 2014; Schwarz et al., 2014). Genomic occupancy studies revealed that OCT4/SOX2/NANOG (OSN) bind overlapping regulatory elements, including their own promoters, and cooperatively regulate a network that self-maintains pluripotency. The OSN core serves as a

platform for recruitment and binding of additional factors such as co-activators, co-repressors, regulatory RNAs or epigenetic factors (Young, 2011). In general, OSN often bind regions containing enhancer activity. Interestingly, gene ontology terms of these regulatory elements reveal enrichments in 'self-renewal' and 'differentiation', further highlighting a role of these core pluripotency factors in generating a bi-stable state (Young, 2011). Although OSN are expressed in both ground state and primed ESCs, their binding differs considerably. Upon transition of serum to 2i, OSN binding is rapidly reconfigured at enhancers increasing at elements with canonical Wnt transcription factor motifs and decreasing at targets of ERK signaling. Rewiring of these core pluripotency factors also correlates with transcriptional activity in ESCs+2i potentially stabilizing the ground state (Galonska et al., 2015b).

Besides the three core pluripotency factors, ESCs express a repertoire of ancillary pluripotency regulators and co-factors, such as KLF2, ESRRB, KLF4, PRDM14, LASS4, TFCP211 and TBX3. Although individual ancillary factors are usually dispensable for ESC integrity, in combination they stabilize each other's expression through several feedback loops reinforcing, fine-tuning and buffering the transcriptional pluripotency network (Dunn et al., 2014). While the core pluripotency factors are expressed in all ESCs, there are substantial differences in the dependency on ancillary factors in ground state and primed ESCs. In general, in ESCs+serum several of these proteins are heterogeneously expressed, reflecting their mixed population of naïve and primed cells. Interestingly, primed ESCs also rely on different subsets of ancillary factors, so-called "modules", highlighting their context-dependent roles (Hackett and Surani, 2014). For instance, the MYC module promotes rapid transition through the G1 phase of the cell cycle and inhibits differentiation in ESCs+serum (Chambers and Tomlinson, 2009). In ground state ESCs on the other hand, the MYC module is almost entirely silenced (Marks et al., 2012) and therefore dispensable for self-renewal. Also ESRRB, a critical component for supporting the pluripotency network in serum ESCs, is not required in ESCs+2i (Grabole et al., 2013). In summary, OCT4, SOX2 and NANOG in concert with several ancillary factors control a cascade of pathways that are interconnected and control pluripotency, self-renewal and cell fate determination (Loh et al., 2006).

1.2 Chromatin and epigenetic modifications in ESCs

Chromatin is the central regulatory unit of genetic material in eukaryotic cells. 147bp of DNA is wrapped around a histone octamer composed of two copies of each core histone proteins: H2A, H2B, H3 and H4 forming a nucleosome, the basic chromatin element. Linker DNA connects adjacent nucleosomes forming a so-called 10nm "beads-on-a-string" fiber (Olins and Olins, 1974). Histone 1 (H1) is not part of the core nucleosome but is capable of binding 20bp of linker DNA interconnecting nucleosomes to form higher order chromatin structures (Ramakrishnan, 1997). Importantly, histones are subjected to numerous post-translational modifications (PTMs) that influence the biophysical properties of the nucleosome (Kouzarides, 2007). In combination with DNA methylation, these histone modifications form the basis of epigenetic inheritance regulating gene function (**Figure 3**).

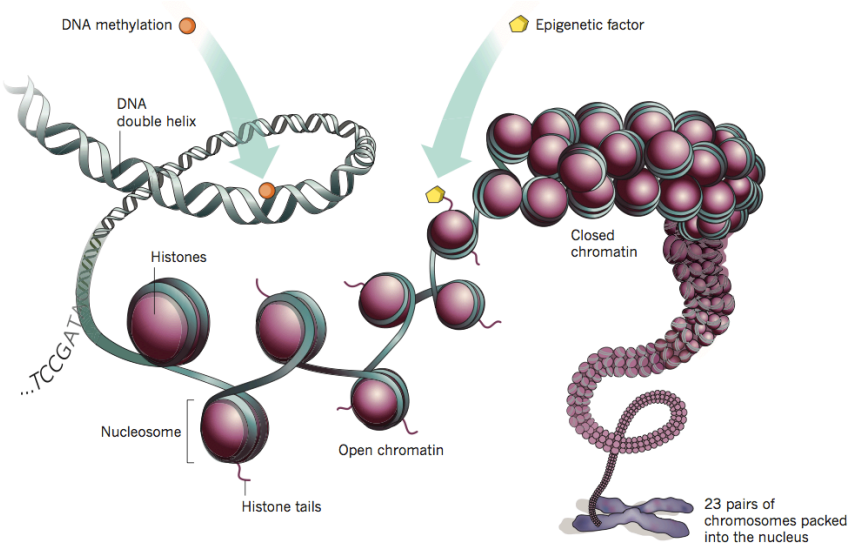


Figure 3. Scheme of chromatin structures. DNA is wrapped around histone octamers giving rise to the nucleosome. Chromatin can then further be organized in higher order structures. Epigenetic marks such as DNA methylation or histone PTMs may alter the transcriptional and compaction state of DNA without changing the underlying DNA sequence. Taken from (Marx, 2012)

They cooperate with nucleosome remodeling complexes and factors regulating chromatin structure to define local and global chromatin dynamics enabling specific genomic regions to be expressed or silent. In general, chromatin can be functionally and structurally classified into "euchromatin" and "heterochromatin". Euchromatin appears as structurally loose and accounts for most transcriptional activity within the cell. Heterochromatin by contrast, is lowly transcribed, forms highly compact

structures and is characterized by repressive histone modifications (Felsenfeld and Groudine, 2003).

1.2.1 Chromatin organization of ESCs

Through self-renewal, ESCs retain the potential to give rise to any cell type of the embryonic tissue. This ability requires a high grade of genome plasticity allowing the ESCs to enter any differentiation path for following lineage-specification. The chromatin of ESCs has several distinct features that allow preservation of the developmentally permissive state. In this regard, ESC chromatin is less condensed and is therefore classified to be in an "open" conformation. This observation is also reflected by an increased ratio between euchromatin and heterochromatin, which is significantly higher compared to differentiating cells (Bhattacharya et al., 2009; Meshorer and Misteli, 2006). Electron microscopy allowed detailed and high-resolution visualization of chromatin structures. Indeed, chromatin in ESCs is evenly spread and largely devoid of heterochromatic regions. In differentiated cells by contrast, chromatin appeared more heterogeneous with distinct blocks of compaction, particularly at the nuclear periphery and around the nucleolus (**Figure 4**) (Efroni et al., 2008; Park et al., 2004; Savic et al., 2014). Importantly, this ESC-specific pattern of evenly spread chromatin organization has also been observed in the ICM epiblast cells of the developing embryo, excluding an artifact of long-term ESC culturing (Ahmed et al., 2010). Recently, super-resolution live cell imaging shed further light on the detailed chromatin dynamics in ESCs and differentiated cells. In line with an open state, the dispersed chromatin domains in ESCs are highly dynamic, especially in central nuclear positions. Upon early differentiation, this flexibility decreases in concert with an increase of compact and static DAPI-dense structures (Nozaki et al., 2017). There is further evidence that support the concept of an open chromatin state in ESCs. Fluorescence recovery after photobleaching (FRAP) experiments have suggested that ESC chromatin contains several loosely bound architectural proteins such as linker histones and heterochromatin protein 1 (HP1). In differentiated cells however, this dynamic fraction was not detected (Meshorer et al., 2006). Recently, also considerable differences in chromatin organization between ground state and primed ESCs were observed. Ultra-high resolution microscopy by STORM investigating chromatin structures at the nucleosome level revealed that ESCs+2i

have a more open chromatin configuration when compared to ESCs+serum. So-called "nucleosome clutches" are significantly smaller in size and contain less nucleosomes per clutch in ESCs+2i. Taken together, pluripotent stem cells keep a largely open and dynamic chromatin conformation allowing the maintenance of a transcriptionally plastic state, which is essential for directing the different transcriptional programs during following lineage specification. Upon ESC differentiation, large-scale genome silencing takes place and ESC chromatin undergoes structural remodeling toward a highly condensed heterochromatic and transcriptionally repressed form (Bhattacharya et al., 2009; Meshorer and Misteli, 2006).

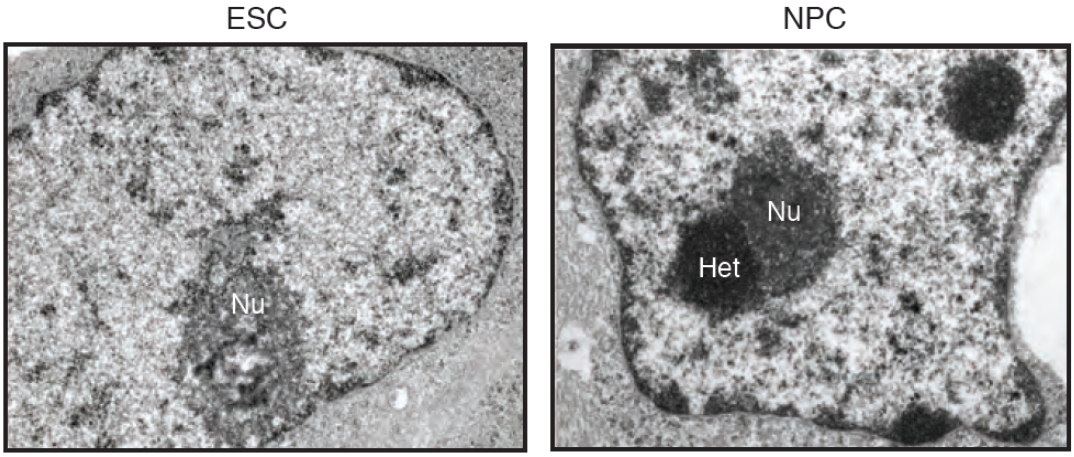


Figure 4. Nuclear re-organization during differentiation. Electron microscopy images from ESC and NPC nuclei. The chromatin in ESCs is in a largely open conformation. Upon differentiation compact heterochromatin blocks form around the nucleolus and at the nuclear periphery. Nucleolus (Nu), Heterochromatin (Het). Taken from (Savic et al., 2014)

1.2.2 Histone modifications

Histones proteins contain a globular motif that allows protein dimerization and assembly of the octamers complex forming the basis of the nucleosome. Their N-terminal domains however extend into an unstructured tail that protrude from the octamer and are able to form additional contacts with DNA and chromatin associated proteins. A striking feature of histones – particularly of their tail domains – is the large number and types of posttranslational modifications (PTMs) they are subjected to. Although many more histone PTMs have been reported, most research has focused on the small covalent modifications methylation (me), acetylation (ac) and phosphorylation (ph) (Kouzarides, 2007). Importantly, these epigenetic marks have a large impact on the biophysical properties of the nucleosome and on the surrounding

chromatin. This can occur directly through steric hindrance of histone-histone or histone-DNA interactions by the modification itself or via regulating the recruitment of effector proteins that specifically recognize the histone PTM (Taverna et al., 2007). The large number of different PTMs expands the regulatory properties of histones beyond that of a structural scaffold upon which the DNA is wrapped (Tee and Reinberg, 2014). Deciphering the co-called "histone code" has been a major research focus in the past years. Several histone PTMs have been characterized in relation to their association and impact on the regulation of gene expression (Strahl and Allis, 2000). Thus, H3K4me1/3, H3K9ac, H3K14ac, H3K27ac H3K36me3, H4K20me are considered as "active" histone marks, since they are mainly associated with euchromatic regions containing actively transcribed genes. By contrast, H3K9me2/3, H3K27me2/3, H3K79me3, H2BK5me3 are defined as "repressive" histone modifications associated with transcriptionally silent regions of the genome (Kouzarides, 2007). All these PTMs are reversible and in a dynamic equilibrium mediated by their addition through "writer" and their removal via "eraser" enzymes. Their functional impact is usually mediated by so-called "reader" proteins, which specifically recognize PTM and trigger a biological output (Atlasi and Stunnenberg, 2017). Further complexity in chromatin organization is mediated by the incorporation of histone variants that slightly differ in their amino acid sequence from the canonical histones. These replication-independent histones are linked to specific functions such as transcriptional activation (H3.3), kinetochore assembly (CENPA), DNA repair and recombination (γH2AX) or X chromosome inactivation (macroH2A) (Sarma and Reinberg, 2005). Taken together, histones and their PTMs provide a near endless combinatorial complexity that can function as a regulatory unit involved in all nuclear processes such as transcription, DNA replication, DNA damage response and nuclear architecture. For this reason the concept of "histone code" has been gradually replaced by the more suitable concept of "histone language" (Rothbart and Strahl, 2014).

1.2.2.1 Histone modifications and bivalent marks in ESCs

As for the chromatin structure, the epigenome of ESCs is highly malleable and dynamic enabling these cells to undergo a wide range of lineage specifications. A global lack of constitutive heterochromatin in ESCs is reflected by reduced levels of

the repressive histone marks H3K9me2/3 and H3K27me3 when compared to differentiated cells (Hawkins et al., 2010; Meshorer et al., 2006). Conversely, high levels of several active histone modifications such as H3K4me3, H3K27ac and H4ac that contribute to keeping the chromatin open and accessible are linked to the mainly euchromatic state of the ESC genome (Azuara et al., 2006; Gaspar-Maia et al., 2011a; Liang and Zhang, 2013). Furthermore, the ES cell genome is transcriptionally hyperactive and transcribes normally silenced repetitive elements as well as coding and non-coding regions, resulting in increased levels of total RNA and mRNA (Efroni et al., 2008). Despite their transcriptional permissive state, ESCs must retain a certain grade of gene silencing to maintain pluripotency and genome stability. Particularly genes involved in differentiation processes need to be kept in a repressed state, without however silencing them irreversibly (Reik, 2007). Consequently, these genes are maintained in a transcriptionally poised state, characterized by the co-occurrence of H3K4me3 and H3K27me3 on their promoters. H3K4me3 is considered an "active" mark that is mainly detected at promoters of transcriptionally active genes. Its deposition is catalyzed by a subset of Trithorax group (trxG) proteins such as SET1, MLL1 and MLL2 (Shilatifard, 2012). Polycomb group proteins by contrast, form multi-subunit Polycomb repressive complexes (PRCs) 1 and 2. PRC1 catalyzes ubiquitination of H2A Lys 119 (H2A119ub) and is also directly involved in mediating chromatin compaction. PRC2 is composed of the core factors EZH2, EED and SUZ12 and catalyzes H3K27me3. H3K27me3 is a central mark in the establishment of repressive chromatin and is mainly deposited at CpG-rich promoters and intergenic regions (Voigt et al., 2013; Zhao et al., 2007). The surprising discovery of both an active and a repressive histone PTM on the same promoter was first identified in mouse ESCs and was recently shown to occur on the same nucleosome and same allele by a co-ChIP approach (Weiner et al., 2016; Zhao et al., 2007). According to the current model, this "bivalent" epigenetic signature safeguards the ESCs from unscheduled perturbations in gene expression (Atlasi and Stunnenberg, 2017). In this regard, H3K4me3 is thought to facilitate recruitment RNA polymerase II and prevents permanent silencing, while H3K27me3 assures that gene expression is kept at low levels (**Figure 5**). Accordingly, most bivalent genes are marked by RNA Pol II occupies in its paused form phosphorylated at Serine 5 (S5p) on the C-terminal domain. Concurrently, Ser 2 (S2p)- phosphorylated RNA Pol II, a hallmark of productive mRNA elongation, is absent at these genes (Brookes et al., 2012). Upon

differentiation toward the neuronal lineage, several bivalent genes loose the H3K27me3 mark and became actively expressed. Genes that need to be entirely silenced by contrast, retain H3K27me3 coinciding with the complete removal of H3K4me3 (**Figure 5**) (Bernstein et al., 2005; Bernstein et al., 2006). These findings led to the intriguing model that bivalent domains maintain developmental genes in a silent state in ES cells while keeping them poised for subsequent activation upon differentiation (Voigt et al., 2013). A functional role for bivalent marks in ESCs was determined by the genetic depletion experiments of the core PRC2 components EED and SUZ12, which causes an entire loss of H3K27me3. These ESCs lacking H3K27me3 were only marginally impaired in their self-renewal potential but showed elevated expression of lineage-specific genes (Azuara et al., 2006; Pasini et al., 2007). In contrast to the mild effects on viability and self-renewal, all H3K27me3-deficient ESCs display a largely impaired differentiation potential (Leeb et al., 2010; Pasini et al., 2007). This phenotype is also recapitulated by a defect in EpiSCs viability and self-renewal in the absence of H3K27me3 and parallels the post-implantation lethality observed in PRC2 KO mice during gastrulation (Faust et al., 1995; Pasini et al., 2004). Detailed investigations of promoter bivalency among the multiple pluripotency states revealed striking differences, largely due to altered levels or distribution of H3K27me3. H3K4me3 on the other hand, seems to be less dynamic and was shown to remain relatively stable in different ESC states and during human ESC differentiation (Hawkins et al., 2010; Marks et al., 2012). Promoter bivalency was originally reported in primed ESCs. Thus, most reports investigating its functional role were performed in ESCs+serum. In 2012 however, bivalency was examined in ESCs+2i that represent a more homogenous population and show a tighter control of lineage-associated gene expression (Marks et al., 2012). Surprisingly, these ground state ESCs had a reduced occupancy of H3K27me3 at developmental promoters and were consequently classified to have a reduced number of bivalent genes (Marks et al., 2012). Nevertheless, it should be mentioned that the average H3K27me3 still is enriched at promoters of these genes in ground state ESCs but were simply not assigned as bivalent because they fell below a arbitrary threshold previously used to define H3K27me3 peaks (Voigt et al., 2013). Consistent with these observations, low levels of promoter bivalency were found in pre-implantation embryos but sharply increased post-implantation. Taken together, bivalent domains keep developmental genes in a primed state that allows either their rapid activation or stable silencing upon

differentiation. This epigenetic feature also provides an elegant mechanism to explain the plasticity of the ESC genome (Voigt et al., 2013). Although bivalency was originally assumed to be stem cell-specific, this epigenetic feature has also been identified in terminally differentiated cells such as dopaminergic neurons, though at lower frequencies compared to ESCs (Ferrai et al., 2017). This suggests that bivalency might provide a general mechanism for dynamic responsiveness to signals such as environmental cues, also in somatic cells (Jadhav et al., 2016). Further genome-wide mapping studies of histone PTMs in ESCs have associated distinct histone marks with certain regulatory elements, such as H3K4me1 and H3K27ac with active enhancers or H3K4me3 and H3K27me3 with active and repressed promoters, respectively (Calo and Wysocka, 2013; Zhou et al., 2011).

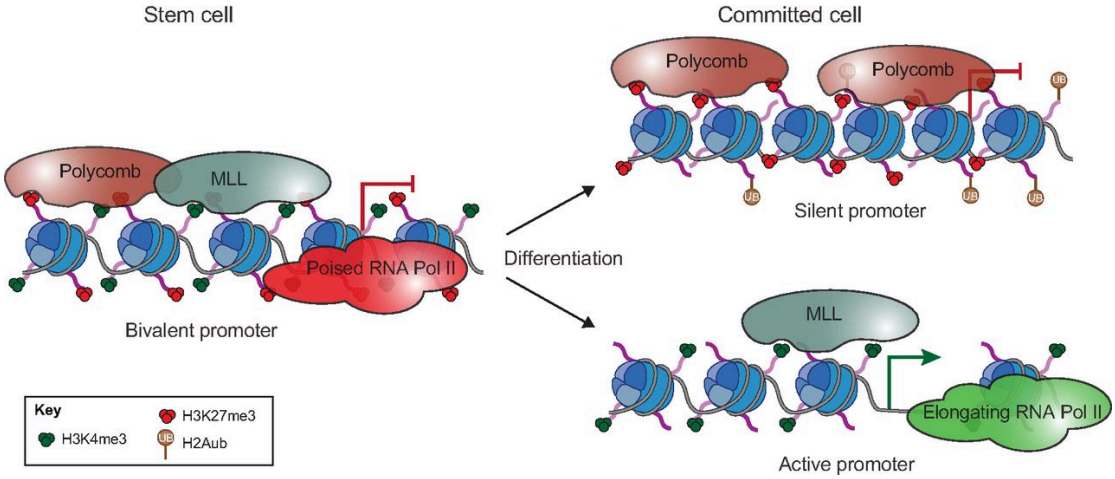


Figure 5. The role of bivalent genes in ESCs and upon differentiation. Promoters of developmental genes are in a bivalent state in ESCs characterized by the presence of both H3K4me3 and H3K27me3. Together, these epigenetic marks lead to the recruitment of RNA Pol II in poised state. Upon differentiation, one of the two bivalent marks is usually lost leading to the permanent activation or silencing of the respective gene depending on gene function and cell type. Taken and modified from (Aloia et al., 2013)

1.2.3 DNA methylation

DNA methylation is among the best-studied epigenetic modification and it is conserved among most plant, animal and fungal models. Since it does not alter the Watson-Crick base-pairing, DNA methylation can be considered as a classical epigenetic modification. The DNA is typically methylated in the CpGs dinucleotide sequence context on carbon 5 of the cytosine base (5mC) (**Figure 6**). Methylation on both opposing DNA strands within the palindromic sequence allows successful maintenance of the epigenetic mark during DNA synthesis. Interestingly, its impact on transcription largely depends on the methylation position within the transcriptional

unit. Associated to promoters, DNA methylation blocks initiation and hence, acts as a transcriptional repressor. Within gene bodies however, 5mC correlates with active transcription and might even stimulate transcriptional elongation (Jones, 2012). Globally, most CpGs remain methylated at an average of approximately 60-80% in mammalian cells (Smith and Meissner, 2013). Highest enrichments of DNA methylation are observed at repetitive elements, satellite DNA, retrotransposons and in gene bodies (Ehrlich et al., 1982; Li and Zhang, 2014). Although globally depleted in the mammalian genome, CpG dinucleotides are found in clusters of so-called CpG islands (CGIs). CGIs are often associated with promoters of housekeeping and developmental regulator genes and are largely resistant to DNA methylation (Deaton and Bird, 2011). Thus, expression of these genes is typically regulated by transcription factors, while their repression can be mediated by various mechanisms such as H3K27me3 (Gal-Yam et al., 2006; Taberlay et al., 2011). Methylation at promoter-associated CGIs is rare but occurs at genes that are subjected to long-term and stable repression. Imprinted genes, promoters of genes exclusively expressed in germ cells and genes located on the inactive X chromosome represent such examples. In general there are two mechanisms by which DNA methylation my influence transcriptional silencing. The methyl marks can either directly or indirectly block transcription factor binding through steric interference or via recruitment of specific methyl-binding proteins, respectively (Tate and Bird, 1993). In the latter case, “reader” proteins that contain methyl-CpG binding domain (MBD) specifically recognize DNA methylation. These factors often recruit further effector proteins mediating transcriptional repression (Fatemi and Wade, 2006). For instance, the methyl-CpG binding protein 2 (MeCP2) binds methylated DNA sequences and recruits HDACs as well as H3K9me3-catalyzing enzymes in order to deposit repressive histone marks and induce transcriptional silencing (Fuks et al., 2003). Alternatively, DNA methylation was proposed to directly impair binding of NRF1 in mouse ESCs (Domcke et al., 2015).

1.2.3.1 DNA methylation machinery

Cytosine methylation is catalyzed the family of DNA methyltransferase (DNMT) enzymes, that include DNMT1, DNMT3A, DNMT3B and DNMT3L (**Figure 6**). DNMT1 recognizes hemimethylated DNA and is involved in maintaining the

methylated state after DNA replication. DNMT3A and DNMT3B associate with the enzymatically inactive DNMT3L and mediate *de novo* DNA methylation acting preferentially on unmethylated DNA (Yokochi and Robertson, 2002). DNMT3C, a novel rodent-specific member of the *de novo* DNMT family, has recently been identified to mediate DNA methylation and silencing of retrotransposon elements during spermatogenesis. Consequently, animals depleted for DNMT3C were sterile, suggesting its specific role for male germ cell development (Barau et al., 2016). How *de novo* DNMTs are recruited to the genome in order to establish correct DNA methylation patterns is not entirely understood.

At least to some extent, their genomic binding is guided by histone modifications. In this regard, H3K4me3 at promoters of active genes has been shown to prevent recruitment of *de novo* DNMTs in order to maintain a hypomethylated state of their CGIs (Otani et al., 2009). In addition, the PWWP domain of DNMT3B was shown to specifically recognize H3K36me3, a histone mark that is co-transcriptionally deposited by SETD2 at gene bodies of active genes. Accordingly, DNMT3B binding is enriched over transcribed regions and its depletion reduces 5mC levels over transcribed gene units (Baubec et al., 2015). A recent study further highlighted a preferential association of the isoform DNMT3A1 with H3K27me3-positive CGIs in ESCs and neuronal progenitors (NPCs). In line with a hypomethylated state of the CGIs, DNMT3A binding did not directly overlap with the H3K27me3 signal but was observed in regions adjacent to it (Manzo et al., 2017). In contrast to the *de novo* DNMTs, recruitment and regulation of the methylation maintenance machinery has been studied extensively. During mitosis, both daughter cells must re-establish correct methylation patterns ensuring the epigenetic inheritance. This mechanism is tightly regulated throughout replication to ensure exclusive targeting of DNMT1 to hemimethylated DNA in order to re-establish symmetrically methylated CpGs. While the DNMT1 transcript is constitutively expressed, DNMT1 protein levels fluctuate throughout the cell cycle in dividing cells, peaking at early S-phase when its activity is most required (Kishikawa et al., 2003). Via its interactions with proliferating cell nuclear antigen (PCNA) and E3 ubiquitin protein ligase UHRF1 (also known as NP95), DNMT1 is recruited to sites of DNA replication. UHRF1 specifically binds hemimethylated DNA through its SET-and RING-associated (SRA) domain correctly directing DNMT1 to the parental, methylated strand (Arita et al., 2008; Chuang et al., 1997; Sharif et al., 2007). Genetic depletion of UHRF1 causes global DNA

hypomethylation and embryonic lethality, largely phenocopying the effects of DNMT1 ablation (Bostick et al., 2007). Furthermore, and in strong contrast to *de novo* DNMTs, DNMT1 is also structurally dependent on a hemimethylated substrate in order to prevent cryptic activity on unmethylated DNA (Song et al., 2011). Besides the core PCNA-UHRF1-DNMT1 interaction, several additional chromatin associated factors and histone modifications have been suggested to regulate DNMT1 recruitment and stability. For instance, the histone acetyltransferase TIP60 and the methyltransferase SET7 modify specific DNMT1 residues and trigger its polyubiquitylation through UHRF1 resulting in DNMT1 targeting for proteasomal degradation (Du et al., 2010; Esteve et al., 2009). This DNMT1-destabilizing pathway is opposed during the peak of DNMT1 activity at early S-phase by histone deacetylase 1 (HDAC1) in complex with DNA-bound DNMT1 (Robertson et al., 2000). Efficient DNMT1 targeting to heterochromatic DNA is assured by the specific binding of UHRF1 to H3K9me2 and H3K9me3, highlighting a connection between DNA methylation and repressive histone marks (Rothbart et al., 2012). Taken together, efficient DNA methylation maintenance is controlled at multiple levels to ensure that DNMT1 activity is stabilized only when in complex with hemimethylated DNA following DNA replication. This provides the necessary fidelity to recreate a precise and global methylation landscape (Smith and Meissner, 2013).

1.2.3.2 DNA demethylation

5mC is a highly stable but reversible epigenetic modification. DNA demethylation is achieved either by passive replication-dependent dilution or by an active and enzyme-dependent removal of the methyl mark. In the case of passive dilution, successive rounds of DNA replication, in which DNMT1 fails to re-establish full methylation, results in daughter cells that have lost the epigenetic mark. Active 5mC removal by contrast, occurs in a stepwise chemical oxidation, which is catalyzed by the methylcytosine dioxygenase ten-eleven translocation (TET) proteins TET1, TET2 and TET3, and ultimately leads to removal of the modification (**Figure 6**) (Wu and Zhang, 2014). In this process 5mC is converted into 5-hydroxymethilcytosine (5hmC), which can be further oxidized to 5-formylcytosine (5fC) and finally into 5-carboxylcytosine (5caC). Both 5fC and 5caC can be recognized and removed by thymine DNA glycosylase (TDG) in the context of the base excision repair (BER)

machinery ultimately restoring an unmethylated C:G base pair (**Figure 6**) (He et al., 2011). Genome-wide binding analyses revealed prominent association of TET1 at CGIs of housekeeping and developmental gene promoters as well as at enhancer elements (Williams et al., 2011). Accordingly, 5hmC – the best-studied intermediate of the cytosine oxidation cascade – is mainly found at enhancers, and depletion of the three TET enzymes in mouse ESCs results in increased DNA methylation at 15–25% of all enhancers (Lu et al., 2014).

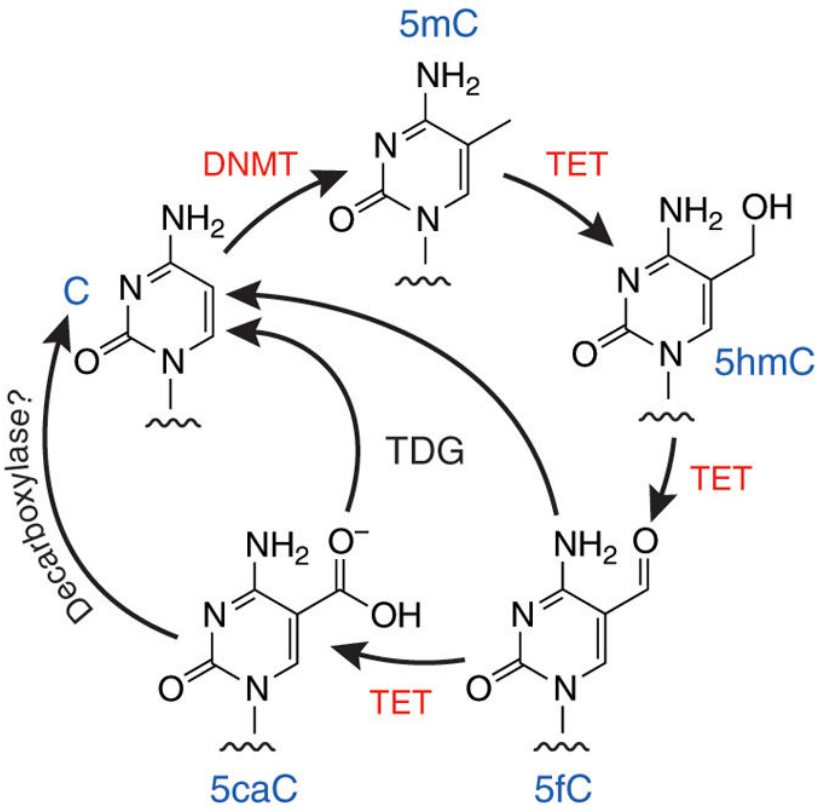


Figure 6. The DNA methylation cycle. DNA methylation is mediated by DNMTs at carbon 5 of the cytosine base (5mC). DNA demethylation occurs in a stepwise chemical oxidation to 5hmC, 5fC and 5caC catalyzed by TET enzymes. In the context of the base excision repair, TDG is able to restore the cytosine to an unmethylated state. Taken and modified from (Huang and Rao, 2012)

1.2.3.3 DNA methylation reprogramming *in vivo*

Although average genome-wide DNA methylation is kept high and bulk 5mC patterns are static across most tissues, two major rounds of rapid demethylation take place during germ line and pre-implantation development. Shortly after fertilization, the hypermethylated paternal genome undergoes a rapid, almost complete loss of methylation (Smith and Meissner, 2013). Mechanistically, this demethylation of the sperm genome was initially explained by TET3-driven oxidation of 5mC to 5hmC.

Recent studies however showed that depletion of TET3 prevents 5hmC formation but does not affect the early demethylation of the paternal genome (Amouroux et al., 2016). Thus, an alternative – and so far unknown – mechanism for DNA demethylation in the early mouse zygote was suggested (Amouroux et al., 2016). After this initial pulse, global methylation levels are continuously decreased reaching a minimum in the pre-implantation blastocyst stage. Different than the rapid demethylation, this second wave of demethylation is TET3-mediated hydroxymethylation dependent and is thought to counteract zygotic DNMT3A and DNMT1 activities (Amouroux et al., 2016). In contrast to the paternal genome, maternal DNA methylation is lost during early embryonic development through passive replication-dependent dilution (Wu and Zhang, 2014). In general, this post-fertilization demethylation is believed to be important for epigenetic reprogramming of the oocyte and sperm genome enabling them to enter the totipotent state (Burton and Torres-Padilla, 2014). Upon exit of pluripotency, levels of *de novo* DNMTs markedly increase leading to a sharp genome-wide re-methylation in the post-implantation epiblast (Smith and Meissner, 2013). Interestingly, a second round of DNA demethylation takes place in post-implantation primordial germ cells (PGCs) and is closely linked to PGC reprogramming for induction of the pluripotency program (Atlasi and Stunnenberg, 2017).

1.2.3.4 DNA methylation in ESCs

DNA methylation reaches its lowest levels in pre-implantation embryos and PGCs. This global 5mC erasure is thought important to remove epigenetic barriers enabling the pluripotent state of these cells (Hackett and Surani, 2013). Despite deriving from the ICM and preserving a naïve pluripotent capacity, ESCs+serum accumulate high levels of global DNA methylation, which is usually associated with primed or lineage-restricted cells (Meissner et al., 2008). By contrast, ground state ESCs+2i have been shown to retain a largely hypomethylated genome with global 5mC levels three-fold lower relative to ESCs+serum (Ficz et al., 2013; Habibi et al., 2013). While levels and distribution of 5mC in ESCs+2i are largely comparable to the *in vivo* ICM epiblast cells, ESCs+serum resemble the hypermethylated state of primed post-implantation cells. Strikingly, the most naïve, REX1-positive subpopulation within the heterogeneous ESCs+serum still retain high 5mC levels, underscoring that these

cells might be functionally naïve but not equivalent to ground state ESCs+2i (Ficz et al., 2013; Habibi et al., 2013). Interestingly, 5mC appears to be highly plastic in these ESCs, since transitions from serum to 2i culturing conditions and its reciprocal switch rapidly result in global DNA hypo- and hypermethylation, respectively (Habibi et al., 2013; Shipony et al., 2014). Nevertheless, several loci such as genomic imprints and IAP elements remain resistant to demethylation, which is consistent with their escape from reprogramming during pre-implantation development (Habibi et al., 2013). Mechanistically, it has recently been proposed that impaired DNA methylation maintenance is primarily responsible for the genome-wide demethylation occurring in ground state ESCs (von Meyenn et al., 2016). Their results suggested that reduced levels of UHRF1 and H3K9me2 in ESCs+2i impair recruitment of DNMT1 resulting in replication-dependent dilution of DNA methylation (von Meyenn et al., 2016). Reduced expression of the *de novo* DNMT3A, DNMT3B and DNMT3L in ESCs+2i further stabilize the hypomethylated genome of ground state ESCs (Habibi et al., 2013). Depletion of TET proteins by contrast, did not affect 2i-induced hypomethylation, implying that active demethylation is dispensable for this process (von Meyenn et al., 2016). An in vivo mechanism for maintaining the hypomethylated state in the ICM has recently been described (Graf et al., 2017). Pramel7 is highly expressed in the ICM and essential for pre-implantation development. Molecular characterization revealed that Pramel7 specifically targets UHRF1 for proteasomal degradation through ubiquitination via Cullin 2 (Graf et al., 2017). Strikingly, forced expression of Pramel7 in ESCs+serum causes a loss in global 5mC due to the impaired in DNA methylation maintenance pathway. As a result, Pramel7-overexpressing ESCs+serum display a gene expression signature that clusters even closer to the pre-implantation epiblast cells than ESCs+2i (Graf et al., 2017). Furthermore, these cells are locked in a pluripotent state, since they are unable to undergo differentiation. However, DNA hypomethylation, induced by genetic depletion of UHRF1 or all DNMTs, did not recapitulate the ground state-specific gene expression signature observed in Pramel7-overexpressing ESCs (Ficz et al., 2013; Fouse et al., 2008; Graf et al., 2017; Sharif et al., 2016). This suggests that DNA hypomethylation alone is not sufficient to induce ground state pluripotency and that Pramel7 functions through additional, DNA methylation-independent mechanisms to promote the ground state (Graf et al., 2017). Taken together, DNA methylation shows

a dynamic pattern during early mammalian development that can largely be recapitulated *in vitro* in ESCs+2i or Pramel7-overexpressing ESCs and ESCs+serum.

1.2.3.5 5mC dynamics in development and disease

Catalytic active DNMTs are essential for correct human and mouse development. Genetic depletion of each subunit is lethal, though at different stages of development. While DNMT1 and DNMT3B KOs are embryonically lethal in both species, mice lacking DNMT3A are viable up to 4 weeks after birth (Li et al., 1992; Okano et al., 1999). Although essential for cellular commitment, DNA methylation is not required for the establishment or maintenance of pluripotency. ESCs that lack any traces of DNA methylation mediated by genetic depletion of all three DNA methyltransferases, also referred to as triple knockout (TKO), remain viable and show no notable aneuploidy. Neither self-renewal nor the molecular signature of pluripotency is affected in these cells (Tsumura et al., 2006). ESCs that specifically lack individual DNMTs have also been reported. Loss of DNMT1 results in a rapid reduction of 5mC to a remaining global level of 20%. ESCs KO for both *de novo* DNMT3A and DNMT3B progressively lose nearly all DNA methylation over many cell division rounds, suggesting that DNMT1 alone is not sufficient for ensuring DNA methylation inheritance (Chen et al., 2003; Smith and Meissner, 2013). Similar to TKOs, molecular stem cell identity is not impaired; whereas differentiation is almost completely inhibited in these ES cell lines. Methylation-free ESCs do not upregulate germ-layer-associated markers and fail to silence pluripotency factors. Interestingly, in somatic cells, mutations affecting human DNMT genes have been linked to several diseases such as acute myeloid leukemia (AML) and immunodeficiency and centromere instability (Shah and Licht, 2011). As for the DNMT family, TET proteins are dispensable for self-renewal despite the complete erasure of 5hmC. TET KO ESCs show differentiation defects and are unable to contribute to chimeric embryos, suggesting essential roles of TET proteins for differentiation (Smith and Meissner, 2013). Taken together, both DNA methylation and TET function are dispensable for self-renewal and ESC integrity but fulfill essential functions upon exit from pluripotency and are required for proper lineage specification.

1.2.4 Chromatin remodeling complexes in pluripotency

Nucleosomes function as important structural building blocks for global packaging and organization of genetic information in the nucleus. At the same time they also represent a steric obstacle for any process that requires DNA access, such as gene transcription. By occupying key regulatory DNA elements, nucleosomes can prevent binding of transcription factors and therefore inhibit expression at the level of transcription initiation. Thus, DNA regions close to transcription start sites of actively expressed genes are usually depleted of nucleosomes (Petesch and Lis, 2012). However, spontaneous remodeling of the nucleosome is usually counteracted by its biophysical properties. 14 histone-DNA contacts that are normally present in the nucleosome lead to substantial positional stability and provide an energetic barrier for remodeling. ATP-dependent chromatin remodeling complexes use the energy from ATP hydrolysis for translocation, eviction or exchange (also referred to as repositioning, ejection and editing) of nucleosomes along the DNA template (Clapier et al., 2017). Chromatin remodeling enzymes are often part of large macromolecular complexes. Usually, a single ATP-hydrolyzing factor provides the energy for remodeling, while several additional factors of the complex modulate the ATPase activity and provide specificity to genome-wide binding. Thus, combinatorial assembly of the individual factors provides the required diversity in chromatin remodeling activities to ensure correct gene expression patterns in a cell- and tissue-specific manner (Ho and Crabtree, 2010). By precisely controlling nucleosome positioning and mobility, these complexes are able to promote the transcription of a set of genes, while simultaneously inhibiting inappropriate expression of another set supporting cell-specific identities. Accordingly, varying subunit compositions have been identified in different cell types and tissues during development (Hota and Bruneau, 2016). All ATP-dependent chromatin remodeling complexes contain an ATPase subunit belonging to the SNF2 family of DNA helicases. Based on the similarities and differences in their catalytic ATPases domains, chromatin remodeling complexes are further categorized into four subfamilies: imitation switch (ISWI), chromodomain helicase DNA-binding (CHD), switch/ sucrose non-fermentable (SWI/SNF) and INO80 (**Figure 7**). At least one member of each chromatin remodeling complex subfamily is essential for mouse embryogenesis, highlighting their central role in mammalian development (Chen and Dent, 2014; Gaspar-Maia et al., 2011a).

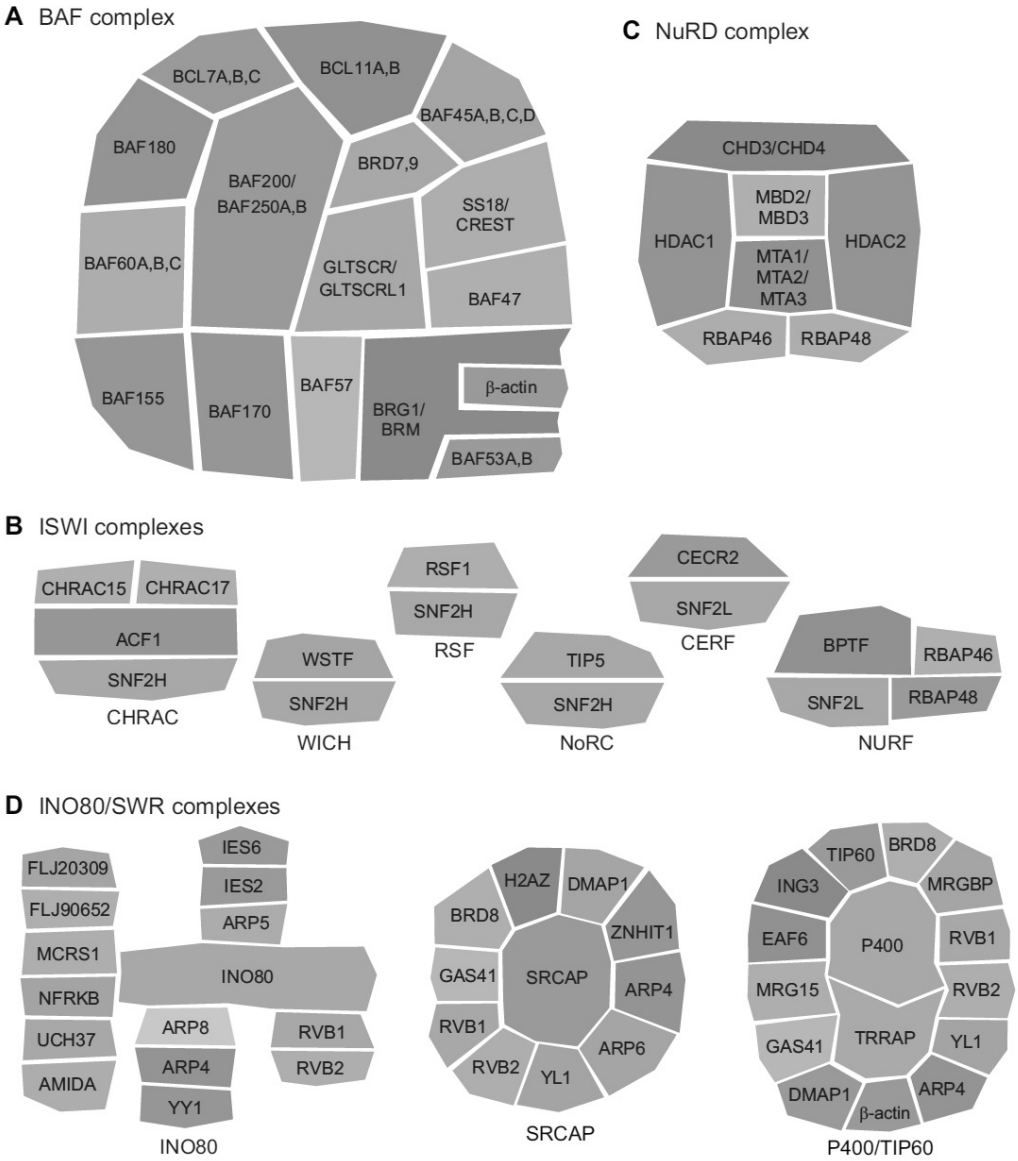


Figure 7. The composition of chromatin remodeling complexes. The subunit composition of some mammalian chromatinremodeling complexes is shown: (A) BAF complex, (B) ISWI complexes, (C) NuRD complex composed of CHD proteins and (D) INO80/SWR complexes. Taken and modified from (Hota and Bruneau, 2016)

1.2.4.1 Roles of SWI/SNF family members in ESCs and during development

The SWI/SNF contains two major complexes: BAF (BRG/BRM-associated factor) and PBAF (polybromo BAF). BAF members consist of at least 15 different subunits and form large, up to 1.5 MDa complexes, that vary in their composition between different cell types and tissues (**Figure 7**) (Lessard and Crabtree, 2010). BRG1 (or SMARC4), its central component containing the ATPase activity, is essential for mammalian development. It was shown to have crucial roles during zygotic genome

activation (ZGA) regulating global H3K4me2. Accordingly, BRG1 KO mice die during pre-implantation development (Bultman et al., 2000). Similar to the *in vivo* development, BRG1 has been shown to be essential in ESCs (Kidder et al., 2009). Interestingly, ESCs have a particular subunit composition that was termed as esBAF, which contains BRG1, BAF155 and BAF60A but lacks BRM, BAF170 and BAF60C (Ho et al., 2009). ESCs depleted for several esBAF components show aberrant self-renewal and an impaired differentiation potential (Hota and Bruneau, 2016; Kidder et al., 2009). Mechanistically, esBAF was shown to function on several levels for maintaining ESC pluripotency by modulating LIF/STAT3 signaling and PRC-mediated repression (Ho et al., 2011). BRG1 binds to promoters of key pluripotency genes and activates their expression. At the same time, esBAF also binds promoters of PRC components and their target genes preventing premature expression of lineage-associated genes (Ho et al., 2009; Kidder et al., 2009). In summary, esBAF is required for early mammalian development and maintenance of the pluripotency transcriptional network in ESCs. However, its function is not restricted to pre-implantation development. BAF chromatin remodeling complexes are also necessary for the development of several further organs and cell types, such as brain, skin, heart, muscle or immune cells (Hota and Bruneau, 2016). Interestingly, in these post-implantation developmental processes, BAF complexes repress expression of core pluripotency factors, implying that its function is highly context- and cell type-specific (Hota and Bruneau, 2016).

1.2.4.2 Roles of ISWI complexes in ESCs and during development

Compared to BAF complexes, ISWI are usually smaller in size. Typically, they are composed of 2-4 accessory subunits that form a complex with one of the two ATPase components SNF2H (SMARCA5) or SNF2L (SMARCA1) (**Figure 7**) (Lazzaro and Picketts, 2001). While SNF2L expression is restricted to the brain and post-natal reproductive tissues, SNF2H is ubiquitously expressed and essential for early embryonic development (Lazzaro and Picketts, 2001; Stopka and Skoultchi, 2003). In the fertilized zygote, SNF2H co-localizes with BRG1 and regulates the initial wave of transcription via its recruitment by transcription intermediary factor 1α (TIF1α) (Torres-Padilla and Zernicka-Goetz, 2006). Thus, its genetic depletion in mice leads to lethality during the pre-implantation development caused by a failure in

proliferation of both trophectoderm and ICM (Stopka and Skoultchi, 2003). After implantation, both SNF2H and SNF2L-containing complexes regulate ectoderm-derived lineages. SNF2H is implicated in promoting proliferation of neural progenitors (Alvarez-Saavedra et al., 2014). Thus, conditional SNF2H mutants display impaired proliferation of granule neuron progenitors and Purkinje cells leading to post-natal neural maturation defects. Mechanistically, SNF2H acts on the C-terminal tail of H2A blocking H1 linker histone loading on chromatin and therefore maintaining the chromatin in a more open conformation (Alvarez-Saavedra et al., 2014). SNF2L by contrast, represses proliferation of neuronal progenitor cells and thereby regulates proper brain size (Yip et al., 2012). Thus, SNF2H and SNF2L form distinct ISWI chromatin remodeling complexes that appear to have antagonistic roles during early brain development. In addition, they also appear to differ in their modes of action. While SNF2H promotes an open and permissive chromatin structure, SNF2L complexes generate a more closed and restricted chromatin configuration. In general, ISWI complexes containing SNF2H seem to have crucial roles during early embryonic development and in progenitor cells, whereas SNF2L-containing complexes are required during differentiation and maturation (Hota and Bruneau, 2016).

One particular ISWI chromatin remodeling complex that was recently shown to fulfill essential roles upon exit of pluripotency is the nucleolar remodeling complex (NoRC) (Leone et al., 2017; Savic et al., 2014). NoRC is composed of two subunits, the ATP-dependent chromatin remodeler SNF2H/SMARCA5 and TIP5, a >200 kDa protein that shares sequence homology with the largest subunits of other members of SNF2H/ISWI-containing remodeling complexes (Erdel and Rippe, 2011; Strohner et al., 2001). Interestingly, NoRC undergoes a switch-of-function upon differentiation. While abundantly present in the nucleoplasm of ESCs, NoRC is specifically recruited to ribosomal RNA genes located in nucleoli only upon exit of pluripotency (Leone et al., 2017; Savic et al., 2014). Once associated with rRNA genes, NoRC mediates their transcriptional silencing via recruitment of DNMTs and histone modifying complexes (Guetg et al., 2012; Santoro et al., 2002). Importantly, this specific tethering mechanism requires a processing step of a long non-coding RNA, which only occurs upon differentiation but not in ESCs (Leone et al., 2017; Savic et al., 2014). Artificial recruitment of NoRC to the rRNA genes could therefore be achieved by solely introducing the processed form of the long non-coding RNA (lncRNA) termed pRNA

(Leone et al., 2017; Savic et al., 2014). Strikingly, this resulted not only in the formation of repressive marks at the rDNA locus but also led to a global spreading of compact heterochromatic structures throughout the nucleus, concomitant with a loss of pluripotency (Savic et al., 2014). Taken together, NoRC serves as an excellent mechanistic example how the site-specific activity of a particular chromatin remodeling complex needs to be kept under tight control for maintaining ESC integrity.

1.2.4.3 Roles of CHD family members in ESCs and during development

In total, nine different chromodomain helicase DNA-binding proteins have been identified (CHD1-9). They largely act alone or – in the case of CHD3 and CHD4 – form a complex termed Nucleosome Remodeling Deacetylase (NuRD) (**Figure 7**). CHD proteins fulfill diverse roles during early and post-implantation embryonic development, as well as in terminally differentiated cell lineages (Hota and Bruneau, 2016). During ZGA, CHD1 regulates the expression of several transcription factors such as OCT4, NANOG or CDX2 that is essential for coordinated ICM and trophectoderm differentiation. Thus, its depletion causes post-implantation lethality (Suzuki et al., 2015). Also CHD4 was shown to have redundant roles during pre-implantation differentiation by restricting the expression of inappropriate genes (O'Shaughnessy-Kirwan et al., 2015). In ESCs, CHD1 regulates both mRNA and rRNA transcription by directly interacting with RNA polymerases I and II (Guzman-Ayala et al., 2015). CHD1 binds open regions of the ESC genome and maintains its euchromatin structure. Accordingly, its depletion leads to defects in self-renewal and pluripotency by inducing the accumulation of compact heterochromatic blocks (Gaspar-Maia et al., 2009). The NuRD complex, which is composed of either CHD3 or CHD4 and several associated factors including HDAC1 and HDAC2, has been under major investigation in ESCs. Although largely dispensable in ESCs, NuRD is required for proper ESC differentiation. At the molecular level, CHD4-containg NuRD complexes associate with C-terminal binding protein 2 (CTBP2) to deacetylate H3K27 and recruit PRC2 for inducing H3K27me3-mediated repression upon differentiation (Kim et al., 2015). Taken together, CHD1 and NuRD complexes seem to have essential roles during pre-implantation development and ESC pluripotency. In addition, NuRD was also found important for later development of synapse formation,

where it represses several developmental genes by its deacetylation activity on H3 lysine 9, 14 and 27 (Yamada et al., 2014).

1.2.4.4 Roles of INO80/SWR complexes in ESCs and during development

INO80 and SWR complexes represent the fourth subfamily of ATP-dependent chromatin remodelers (**Figure 7**). They are characterized by a conserved insertion in their ATPase/helicase domain that mediates their interaction with the two helicases RVB1 and RVB2. In mammalian cells three different, multi-subunit complexes are distinguished: The INO80 complex and two SWR-related complexes SRCAP and P400/TIP60 (Hota and Bruneau, 2016). The INO80 complex binds to promoters of ESC master transcription factor genes and promotes their transcription by recruiting the mediator complex and RNA pol II. Thus, INO80 is crucial for ESC self-renewal and pluripotency (Wang et al., 2014). Besides INO80, an RNAi screen also identified the SWR-related P400/TIP60 complex to be essential for ESC pluripotency (Fazzio et al., 2008). In addition to directly promote transcription, both INO80 and SWR complexes are also able to mediate the exchange and deposition of the histone variant H2A.Z. This histone editing activity occurs at active and poised promoters in ESCs and is required for proper ESC differentiation (Subramanian et al., 2013). In summary, all four groups of ATP-dependent chromatin remodelers mediate crucial functions during early mammalian development. These roles are mediated by both ATP-dependent remodeling of nucleosomes and in an ATP-independent manner by their interaction with auxiliary effector proteins such as transcription factors or histone modifying enzymes.

1.3 3D genome organization

Folding of DNA into chromatin and its spatial organization is essential for all biological processes occurring within the nucleus such as transcription, replication, repair and recombination. Regulation of transcription for instance, does not only occur through promoter-proximal regions but can also be modulated by distal regulatory elements. These enhancers are often far apart in linear DNA sequence – up to 1Mb in the case of the *Shh* gene (Lettice et al., 2003) – but can physically associate with their respective target promoters by looping of the chromatin. In a larger scale, spatially isolated genomic regions, such as lamina-associated domains (LADs) or

topological-associated domains (TADs), have been described to form local sub-compartments that can be permissive or repressive for gene transcription. Thus, the 3D genome architecture is highly regulated in mammalian cells allowing gene regulatory networks in a cell- and tissue-specific manner (Bonev and Cavalli, 2016). The first report of spatial chromosome organization goes back go 1885 and was proposed by Rabl and Boveri (Boveri, 1909; Rabl, 1885). According to their model, chromosomes exist in discrete chromosome territories (CT) in interphase nuclei (Boveri, 1909; Rabl, 1885). Strikingly, this territorial model of chromosomal organization has been generally accepted nowadays. Thus, each chromosome resides in a separate region in the nucleus and becomes randomly re-shuffled after each cell division, since no preferential nuclear positioning was attributed to the individual chromosomes (Cremer and Cremer, 2010). Accordingly, the large majority of all 3D chromatin contacts occur within the same chromosome (Cremer and Cremer, 2010). This model was further supported by mathematical polymer modeling of DNA dynamics and kinetics within the nucleus (Dekker and Mirny, 2016). According to their calculations, loci located on different chromosomes are insufficiently close in space in order to re-establish a reproducible and stable physical interaction within the time frame of one cell cycle (Dekker and Mirny, 2016). Only very recently, technical advances allowed detailed investigation of intra-chromosomal organization in mammalian cells genome-wide. Indeed, these experiments doubtlessly confirmed that the vast majority of DNA contacts occur within chromosomes, while stable interactions between chromosomes are extremely rare (Lieberman-Aiden et al., 2009). In general, chromosome conformation capture techniques (3C, 4C, 5C and HiC) rely on proximity ligation of closely associated DNA regions in 3D. Coupled to high throughput sequencing, HiC results generate maps showing the contact frequencies of physical interactions of two DNA loci. Usually, results of these population-based techniques are further investigated by fluorescence *in situ* hybridization (FISH), which allows analysis of DNA loci positioning in individual cells within a population. Although the resolution of fluorescence microscopy has greatly improved, FISH comes at the cost of genome-wide analyses by focusing on a restricted amount of loci at a time (Dixon et al., 2016).

1.3.1 Hierarchical organization of chromatin

Together these new approaches shed light on a highly complex and multi-layer organization pattern of chromosomes within mammalian cells (**Figure 8**). At the basis lie nucleosome-nucleosome interactions that form a 10nm "beads-on-a-string" fiber (Olins and Olins, 1974; Ricci et al., 2015). Nucleosomes cluster into heterogeneous groups of so-called "clutches" that are cell-type dependent and vary in size (Ricci et al., 2015). As a next level of organization, chromatin can form loops bringing two distant DNA elements in close proximity. As one of the key features of vertebrate genomes, cis-regulatory elements can thereby act on their target promoters over large distances of linear DNA. These interactions were further proposed to form active chromatin hubs promoting transcription by high local concentrations of transcription factors and RNA polymerase II (Dixon et al., 2016). However, long-range chromatin contacts are not limited to enhancer-promoter contacts but can also occur between promoters and its respective transcribed region. These "gene-loops" are thought to reinforce the directionality of RNA synthesis from the promoter and promote the re-start of a transcription cycle after its termination (Bonev et al., 2017; Tan-Wong et al., 2012). As one of the most exciting discoveries of recent HiC analyses was the identification of topologically associated domains (TADs) (Dixon et al., 2012). At the level of several hundred kilo bases (kb), chromosomes are segregated into spatially isolated regions. This partitioning into TADs correlates with linear genomic features such as histone modifications, coordinated gene expression patterns and replication timing (Dixon et al., 2012). Surprisingly, TADs were shown to be largely cell-type invariant and even conserved among several species (Dixon et al., 2012; Rao et al., 2014). Thus, TADs are considered as the basic unit of chromosome folding and are suggested to be an important secondary structure in chromosome organization (Dixon et al., 2016). Further chromatin compartmentalization can be observed at a megabase-scale. Large-range interactions between individual TADs occur across entire chromosomes. This largely bimodal interaction pattern led to the definition of A and B compartments, based on their preferential interaction with each other (A compartments interact with other A domains and vice versa) (Lieberman-Aiden et al., 2009). Independent of 3C-based technologies, further studies investigated in defining genomic regions that are spatially isolated within the nucleus. Lamina-associated domains (LADs) are composed of heterochromatin and late-replicating domains that preferentially locate to the nuclear periphery and add an additional layer of chromatin

organization inside the nucleus (Peric-Hupkes et al., 2010). Taken together, these observations led to the conclusion that chromosome structures are formed in a highly organized and hierarchical manner (**Figure 8**).

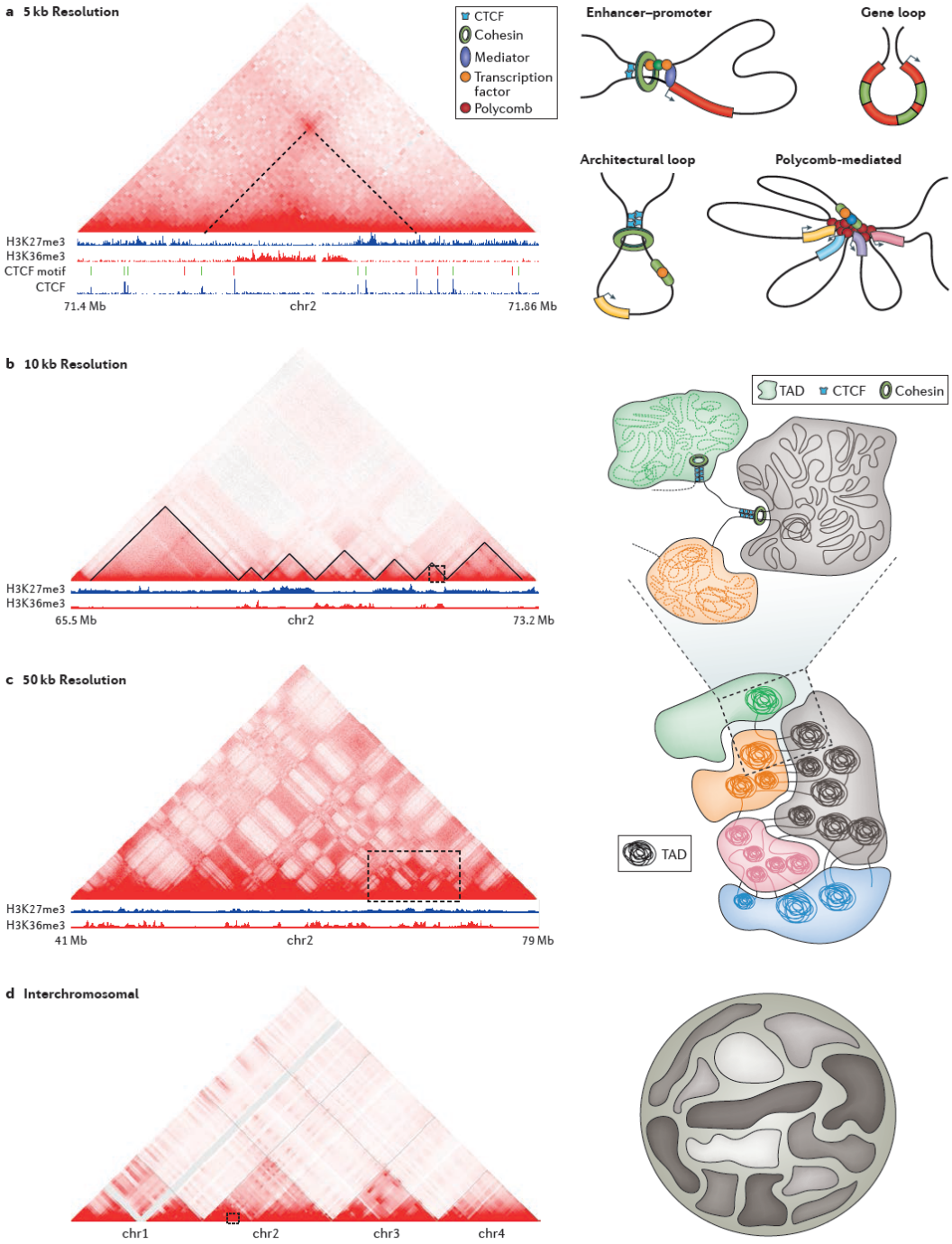


Figure 8. Hierarchical organization of chromatin structures. a. Examples for chromatin loop formation. A loop structure is characterized by the point-shaped peak in local contact frequencies of HiC maps. b. TADs form triangles in HiC maps and are often enriched for loop structures at their boarders. c. Genome partitioning into A/B compartments are highlighted by the checkerboard pattern in low resolution HiC-maps. d. Chromosomes themselves are isolated genomic elements that form chromosome territories consisting each of a single chromosome. Taken from (Bonev et al., 2017)

1.3.2 Chromatin loops

As introduced earlier, chromatin loops bring two distant genomic loci into close spatial proximity (**Figure 8**). This mechanism is mostly acknowledged to bridge enhancers to their respective promoter targets allowing their timely activation and expression. Thus, chromatin looping is a dynamic and cell-type specific process (Bonev and Cavalli, 2016). One well-described example occurs at the locus control region of the β-globin cluster. Via long-range chromatin interactions, it contacts its target genes exclusively in erythroid cells leading to their active expression. Cells from different lineages however, do not form this loop and hence, do not express β-globin genes (Palstra et al., 2003). In order to address which factors are involved in mediating long-distant DNA contacts, a recent study identified CTCF, cohesin, YY1 and ZNF143 enriched at anchors of strong chromatin interaction points (Rao et al., 2014). Indeed, it is now believed that CTCF and cohesin are the main players for the formation of chromatin loops. CCCTC-binding factor (CTCF) binds to the non-palindromic consensus sequence CCGCGNGGNGGCAG, which is mediated by its 11 zinc finger motifs (Ong and Corces, 2014). CTCF was initially identified as an insulator to heterochromatin spreading (Kellum and Schedl, 1991) and is now accepted as one of the main architectural proteins mediating intrachromosomal interactions (Ong and Corces, 2014). Cohesin on the other hand is a ring-shaped complex composed of SMC1, SMC3 and RAD21. Through its structure it can encircle two chromatin fibers and has thereby been shown to hold the two sister chromatids together after DNA replication until the onset of anaphase. Its role as architectural protein forming chromosome loops has only recently been investigated (Ong and Corces, 2014). In ESCs, cohesin associates with the Mediator complex and localizes to active enhancers and is further believed to mediate enhancer-promoter interactions (Kagey et al., 2010). Accordingly, its short-term depletion leads to immediate loss in nearly all loop structures detected by HiC experiments, underscoring its essential role in promoting and stabilizing these 3D structures (Rao et al., 2017). Mechanistically, a model of “loop extrusion” has been proposed to form these long-range contacts mediated by cohesin and CTCF. Although cohesin occupancy is detected at CTCF bound sites in steady-state conditions by ChIPseq, its loading on chromatin is thought to occur elsewhere (Kagey et al., 2010; Zuin et al., 2014). Upon loading, cohesin then travels along chromosome arms by actively pushing DNA through its ring-shaped domain. Consequently, a loop structure will

form and continuously increase in size until the cohesin complex reaches two CTCF-bound boarder elements on each side. This, in turn, traps the cohesin complex and leads to its stable association at these sites, ultimately forming the loop anchor (**Figure 9**) (de Wit et al., 2015; Guo et al., 2015; Nora et al., 2017; Sanborn et al., 2015). Interestingly, the orientation of CTCF binding seems to be crucial for this trapping mechanism, since stable loops are only formed between convergently oriented CTCF sites (Rao et al., 2014).

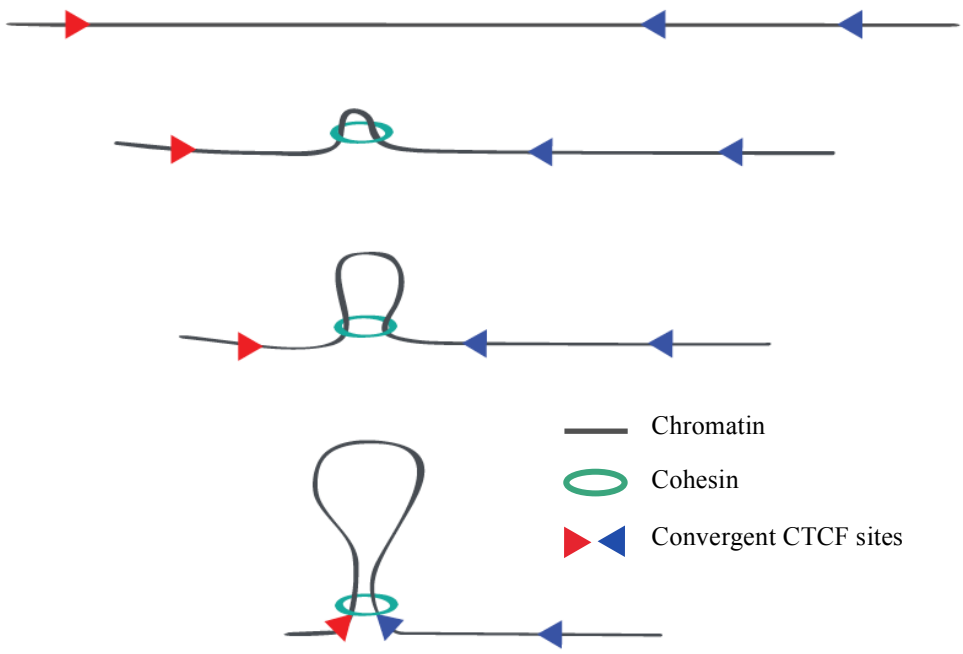


Figure 9. Loop extrusion model. Cohesin loading occurs in a sequence independent mechanism, preferentially at sites of active transcription. By pushing the chromatin through its loop, cohesin travels along the chromosome arms forming a loop structure. At convergently oriented CTCF sites cohesin is trapped resulting in a stable chromatin loop. Taken and modified from (Schwarzer et al., 2017)

Dynamic loading and unloading of cohesin on and from chromatin are essential for this process, since depletion of both loading and unloading factors largely impacted global 3D chromatin structures (Busslinger et al., 2017; Haarhuis et al., 2017; Schwarzer et al., 2017). These studies further highlighted that cohesin association counteracts compartmentalization into A/B compartments. Stabilizing cohesin on chromatin by depleting its releasing factor WAPL led to a global increase of loop structures. This effect seemed to counteracted global genome partitioning into A and B compartments, since these chromosome-wide structures were largely reduced in WAPL KO cells (Haarhuis et al., 2017). Genetic depletion of cohesin loading by

contrast resulted in an increased A/B compartmentalization, suggesting that local and global genome organization mechanisms seem to counteract each other to some extent (Schwarzer et al., 2017). Thus, it has been proposed that cohesin turnover might provide the flexibility necessary for nuclear compartmentalization (Haarhuis et al., 2017). In line with this hypothesis, siRNA-mediated knockdown of cohesin indeed resulted in an increase in chromatin dynamics, as determined by high-resolution live cell imaging (Nozaki et al., 2017).

Although CTCF and cohesin have doubtlessly been shown essential for chromatin looping, they are not the only proteins that are implicated in long-range chromosomal contacts. For instance, the Mediator complex and yin yang 1 (YY1), which interact with cohesin and CTCF, respectively, have been proposed to mediate intrachromosomal contacts in interphase cells (Jonathan A. Beagan, 2017; Kagey et al., 2010). Furthermore, PRC-repressed genes were shown to form ultra-long contacts forming highly dense and compact clusters of repressive chromatin. Strikingly, these contacts were dependent on PRC components, indicating that these epigenetic factors can directly influence the 3D architecture (Boettiger et al., 2016; Joshi et al., 2015; Schoenfelder et al., 2015a).

1.3.3 Topological associated domains

One level above chromatin loops, 3D chromatin structures extend into topological associated domains (TADs). TADs were originally identified by HiC experiments, in which TADs manifest as contiguous square domains along the diagonal of contact frequency maps. This pattern indicates that regions within the same TAD interact with each other more frequently than with regions of adjacent TADs (**Figure 8**) (Bonev and Cavalli, 2016). This implies two basic features of TAD organization: “self-association” within TAD domains and “insulation” between neighboring TADs. These two properties are the basis for the various computational algorithms that have been used for TAD identification. In general, insulation scores of at least 2-fold in interaction frequencies are used as threshold for TAD definition (Dekker and Mirny, 2016). It must therefore be taken into consideration that TADs arise through arbitrarily chosen thresholds from HiC results and are not linked to a specific biological function for their definition. Thus, their calling may alter depending on the depth, quality and resolution of HiC sequencings. Accordingly, TADs were initially

described at a median size of 880kb, whereas subsequent analysis of higher resolution HiC data suggested a smaller median domain size of 185kb (Bonev and Cavalli, 2016). A further degree of inconsistency may be explained by differences in the computational algorithms used to identify them (Dixon et al., 2016). However, the confusion in varying TAD annotations is mainly caused by the fact that TADs – similar to the global 3D chromatin structures – are hierarchical in their organization (Phillips-Cremins et al., 2013). Large TADs can therefore be further divided into smaller domains referred to as sub-TADs. Although TADs and subTADs show many similarities in their structures and insulation degrees, there are some substantial differences among them. Most importantly and in contrast to TADs, subTADs seem to differ between different cell lineages (Phillips-Cremins et al., 2013). In this regard it was proposed that TADs might represent a larger and more invariant level of chromatin organization within which cell-type-specific structures can form to play roles in lineage-specific genome regulation (Dixon et al., 2016).

It is now accepted that the property of insulation is the most determinant for TAD structures. TAD boarders function as boundary elements preventing the ability of an enhancer to activate genes located in different TADs (**Figure 10**) (Hnisz et al., 2016a). In mammals, many of these insulated anchor regions possess strong chromatin loop structures leading to the alternative term of "loop-domain" (Rao et al., 2014). Irrespective of its nomenclature, this observation suggests a strong relationship between chromatin loops and the spatial organization into TADs. Indeed, anchor sites are highly enriched for cohesin and CTCF occupancy. In order to elucidate their causative roles in TAD formation, several recent studies investigated the effect of cohesin and CTCF depletion on TAD structures (Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017). Strikingly, loss of either one of these two factors leads to complete loss in TAD structures, underscoring their essential roles in TAD formation (Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017). Attempts to predict TADs however, has been challenging so far, since only few CTCF-bound sites form stable loops that function as insulating regions. Accordingly, only 15% of all mammalian CTCF-binding sites are located within a boundary, suggesting that CTCF binding alone may be insufficient for the establishment of boundaries.

Further correlation studies revealed a significant overlap of TAD structures with other genomic features such as histone modifications, coordinated gene expression and DNA replication timing (Dixon et al., 2012). Nevertheless, TAD organization is not a

consequence of chromatin marks, since – at least for the Xist locus – TADs remained unchanged in the G9a and EED KO ESCs that lack H3K9me2 and H3K27me3, respectively (Nora et al., 2012). Conversely, TADs did not seem to be determinant for epigenetic landscapes, since neither CTCF nor cohesin depletion largely affected histone modifications genome-wide (Nora et al., 2017; Schwarzer et al., 2017). These results are to some extent contradictory to studies, in which deletion of single boundary elements clearly led to spreading of active chromatin structures into the neighboring TAD (Narendra et al., 2015). Taken together, these findings suggest that TADs do not regulate global chromatin patterns. However, CTCF-occupied boundaries may restrict spreading of active or repressed chromatin in a local and context-dependent manner.

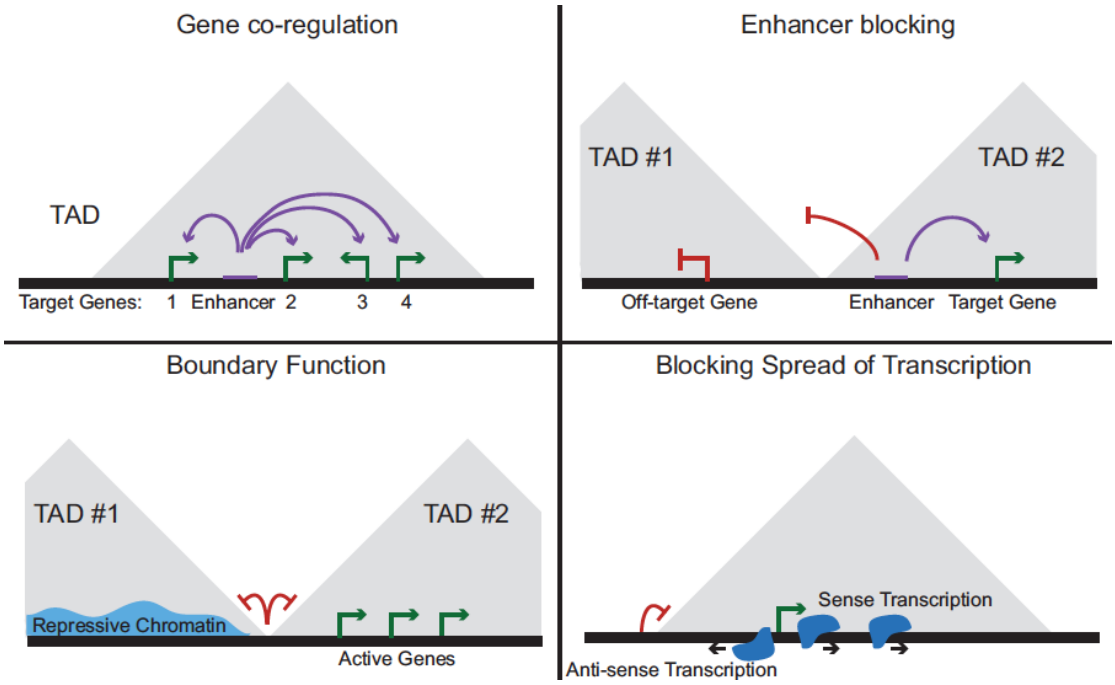


Figure 10. TAD function in gene regulation. Genes within the same TAD are often expressed in a co-regulated manner, presumingly driven by a single enhancer element. TAD boundaries blocks the contact of an enhancer located in one TAD with target genes of another TAD. In a few cases TAD boundaries may also function as insulators for spreading of repressive chromatin into active domains and vice versa. TAD boarders can also function as a physical barrier to antisense transcripts. Taken and modified from **(Dixon et al., 2016)**

As functionally isolated units of the genome, gene expression was shown to occur in a co-regulated manner within TADs (**Figure 10**) (Dixon et al., 2016). Thus, it has been observed that genes within the same TADs are coordinately expressed across different cell types and tissues. This may be particularly important for gene clusters of

functionally related genes, such as for olfactory receptors, which have often been found to reside within the same TAD (Dixon et al., 2016). TAD boundaries on the other hand, seem to fulfill an important insulation function to prevent spreading of gene expression activities (**Figure 10**). Deletion of a TAD boundary on the X chromosome in ESCs for instance, led to increased expression of several genes located on the other side of the boundary (Nora et al., 2012). Strikingly, TAD boarders were even shown to act as physical barriers preventing the linear tracking of molecules along the chromosome. This phenomenon was described by the observation that non-coding transcripts from divergent transcription abruptly ended at TAD boundaries (**Figure 10**) (Austenaa et al., 2015). The most impressive work on elucidating the biological function of TAD boundaries was performed by Mundlos and colleagues (Lupianez et al., 2015). Their investigations revealed that several diseases displaying limb malformation were linked with mutations in a TAD boundary. Consequent rewiring of enhancer-promoter contacts over the lost boundary caused deregulated expression of genes involved in limb bud development. Strikingly, this disease phenotype could be highly recapitulated by introducing deletions, inversions or duplications into the DNA sequence of the respective TAD boundary in mice (Lupianez et al., 2015). There have been further studies, which linked disease phenotypes with disruption of TAD structures. Particularly cancer cells have been proposed to undergo "enhancer hijacking". During this process enhancer contacts are rewired and act on new – potentially proto-oncogene - targets, which can be caused by elimination of TAD boundaries (Hnisz et al., 2016b; Northcott et al., 2014).

1.3.4 Megabase-scaled compartmentalization

Results from HiC experiments revealed an invaluable insight into large-scale spatial organization of chromatin within the nucleus of mammalian cells. They highlighted the segmentation of the linear chromosomes into TADs, which have been shown to be cell type invariant. A further discovery from HiC results was the preferential association of individual TADs with each other over variable genomic distances giving rise to two types of compartments, called A and B (**Figure 11**) (Dixon et al., 2016). This megabase-scaled compartmentalization is visible as a checkerboard pattern of contact frequencies between genomic intervals over the entire chromosome (Schwarzer et al., 2017). While A-compartments highly correlate gene expression and

active histone marks, B-compartments largely represent silent regions of the genome. In contrast to TADs, compartmentalization into A and B are cell-type specific. Thus, TADs themselves are conserved but their preferential association with other TADs can switch through attraction and/or repulsion mechanisms depending on their transcriptional activities. This switching was observed to occur extensively during ESC differentiation or reprogramming of somatic cells to induced pluripotent stem cells (iPSCs) (Bonev et al., 2017; Stadhouders et al., 2018). The mechanism responsible for the formation of A/B compartmentalization remains elusive. However, a recent study highlighted that TADs themselves are not required for A/B compartmentalization (Schwarzer et al., 2017). In fact, deletion of TAD structures by preventing cohesin loading resulted in a more nuanced and finer segregation of the genome into A/B compartments. Importantly, this new compartmentalization directly corresponded to gene expression activities and chromatin modifications of the respective loci (Schwarzer et al., 2017). This observation suggests that chromatin has an intrinsic tendency to self-associate into A/B compartments based on the local epigenetic landscape and transcriptional activity. Furthermore, this megabase-scale of genome organization is – to some extent – opposed by the segregation of the chromatin into TADs, which brings loci of contrasting states together into the same hub (Schwarzer et al., 2017).

1.3.4.1 Nucleolar-associated domains (NADs)

Independent of proximity ligation experiments, many studies have aimed to determine the genomic composition of sub-nuclear compartments, such as the nucleolus and the nuclear lamina (**Figure 11**). Electorn microscopy experiments have shown that nucleolar-associated domains (NADs) contain highly compact heterochromatic structures (Efroni et al., 2008; Park et al., 2004; Savic et al., 2014). However, their identification and characterization has been challenging, since specific targeting of the membrane-less compartment has not been successful so far. To date, attempts to identify NADs have been performed by purification and sequencing of nucleoli (Nemeth et al., 2010; van Koningsbruggen et al., 2010). Although both studies identified mainly inactive regions from all chromosomes associated to nucleoli, their results revealed variable outcomes. While one study observed enriched tRNA genes at NADs (Nemeth et al., 2010), the other report suggested a significant overlap of NADs

with previously published LADs (van Koningsbruggen et al., 2010). Most likely, these discrepancies derive from variations of the sonication-based purification procedure, making a precise definition of NADs impossible. Thus, their comprehensive characterization among different cell types and tissues remains to be determined.

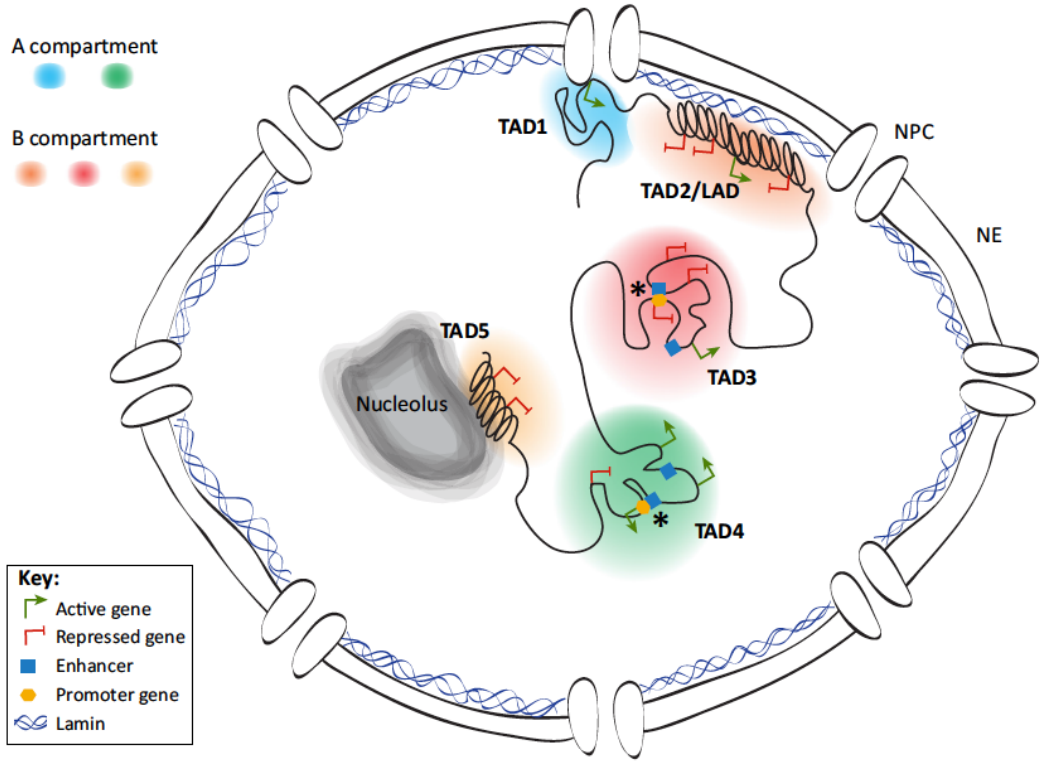


Figure 11. Global genome compartmentalization. Large-scale genome segregation can be observed into A and B compartmentalization of self-associating active and repressed genomic regions, respectively. Further genome organization can be observed by preferential association of mostly heterochromatic regions to the nuclear lamina (LADs) and to the nucleolus (NADs). Taken from (Gasser, 2016)

1.3.4.2 Lamina-associated domains (LADs)

Identification and characterization of lamina-associated domains (LADs) was achieved in 2008 using a method called DNA adenine methyltransferase (dam) identification (DamID). Fusion of the bacterial dam protein to LAMIN B1 mediates site-specific GATC methylation of DNA loci located at the nuclear periphery thereby allowing the identification LADs (Guelen et al., 2008). This and following studies revealed comprehensive maps of LADs in different cell types and species (van Steensel and Belmont, 2017). Mouse and human cells contain 1000 to 1500 LADs, which are between 10kb and 10Mb in size and present on all chromosomes. They

cover approximately 40% of the genome and are generally gene-poor associated with low levels of gene expression. Typically rich in A/T-rich sequences, LADs are also enriched for the repressive histone marks H3K9me2 and H3K9me3 (van Steensel and Belmont, 2017). However, the exact mechanism of anchoring of LADs to the nuclear lamina (NL) in mammalian cells is not fully understood. Results from single-cell DamID experiment revealed that LAD regions contact the NL over continuous stretches, suggesting multivalent interactions that do not include a specific DNA consensus sequence (Kind et al., 2013). At least in part, this interaction is mediated by H3K9me2 and H3K9me3. Deletion or inhibition of the H3K9 methyltransferases G9A, SUV39H1 and SUV39H2 weakened the NL-LAD interaction and its triple KO was sufficient to peel the HBB locus completely away from the NL (Bian et al., 2013; van Steensel and Belmont, 2017). Reciprocal experiments elucidating the role of lamin proteins in mediating NL-LAD interactions revealed surprising results. Deletion of all lamins had no effect on genome-wide contacts with the NL in mouse ESCs, suggesting redundant roles of LAMIN A/C and LAMIN B receptor in generating a scaffold to which LADs are anchored (Amendola and van Steensel, 2015). Experiments investigating gene positioning in single cells revealed that only 30% of LADs contact the lamina at a given time point. Thus, LADs are not constantly associated with the lamina but show some dynamics and seem to randomly attach or detach at every cell cycle (Kind et al., 2013). Although the NL has been accepted as a silent compartment of the genome, investigating its causal role in gene repression has resulted in variable outcomes. Insertion of reporter genes into LADs or their artificial tethering to the NL was reported to lead to a two- to six-fold reduction in expression (Akhtar et al., 2013; Reddy et al., 2008). Other studies by contrast, did not observe any effects on expression or transcriptional activation of their reporters upon NL-localization (Finlan et al., 2008; Kumaran and Spector, 2008). Together, these results suggest that proximity to the NL can cause reduction in gene expression, but only of a subset of genes (van Steensel and Belmont, 2017). Besides generating a map of LADs, DamID experiments also revealed a large proportion of the genome that is preferentially positioned away from the NL. These so-called inter-LADs or ICDs are located in the nuclear interior and account for most transcriptional activity of the genome. Furthermore, they contain high gene densities and are enriched for active histone marks, such as H3K4me1, H3K4me3 and H3K27ac (van Steensel and Belmont, 2017). Correlation studies with HiC experiments revealed a near identical

overlap between inter-LADs/LADs with A/B compartmentalization. While LADs largely correspond to the B compartment, the A compartment overlaps with inter-LADs (Kind et al., 2015). This implies that LADs tend to interact with each other and are spatially separated from active compartments of the genome. Strikingly, there is also a cross talk between LADs and NADs. DamID-mediated labeling of LADs showed that after mitosis NL-associated domains redistributed to both the NL and the nucleolus. This observation indicates that LADs are variably positioned at either the NL or in close association with the nucleolus (Kind et al., 2013).

1.3.5 3D chromatin organization in ESCs and during development

Higher-order chromatin structures in 3D are well known to play essential roles in gene expression regulation. How these structures emerge during development could only very recently be addressed due to technical challenges of low starting material. Three recent studies investigated the organization of the genome during early embryonic development in mice and drosophila using low-cell HiC experiments (Du et al., 2017; Hug et al., 2017; Ke et al., 2017). Interestingly, the genome of mature oocytes at the metaphase II stage completely lacks TADs and exists in a largely unstructured conformation (Ke et al., 2017). In sperm by contrast, extra-long-range interactions (>4Mb) and even inter-chromosomal interactions occur frequently, which is in agreement with a highly compacted sperm genome (Ke et al., 2017). Upon fertilization, large-scale chromatin re-organization of the paternal genome takes place resulting in a largely diminished higher-order structure after fertilization (Du et al., 2017). Subsequent re-establishment of chromatin organization is a slow process and extents throughout the entire pre-implantation development (Du et al., 2017). Importantly, all three studies in parallel reported that this re-organization of the 3D architecture could occur in the absence of transcription, ruling out a major role of the zygotic genome activation for TAD formation (Du et al., 2017; Hug et al., 2017; Ke et al., 2017). By contrast, treatment with aphidicolin abolished TAD formation, attributing an essential role of DNA replication for TAD establishment (Ke et al., 2017). Taken together, these findings highlight that compartmentalization of the genome after fertilization is a slow process, which is not transcription- but replication-dependent. These studies further highlight that pluripotent epiblast stem cells in the ICM have already completely re-established TAD structures. Accordingly, cultured

ESCs show clear genome partitioning into TADs and A/B compartments (Dixon et al., 2012). A next large-scale re-organization of the genome occurs upon exit of pluripotency and during following lineage-specification. Upon differentiation, the open structure of ESC chromatin is remodeled associated with the formation of compact heterochromatic blocks and reduced chromatin dynamics (Bhattacharya et al., 2009; Meshorer and Misteli, 2006; Nozaki et al., 2017). Interestingly, the nucleolus seems to fulfill essential roles during this process serving as an ignition point for global heterochromatin formation and spreading during very early differentiation processes (Savic et al., 2014). Regarding sequence-resolved analyses, DamID and ultra-high resolution HiC have been performed on ESCs and ESC-derived neuronal lineages (Bonev et al., 2017; Peric-Hupkes et al., 2010). As expected, TADs themselves largely remain constant during differentiation to neuronal cells, whereas association between TADs alters globally. Contacts between active TADs become less pronounced, while inactive TADs interact more strongly. Furthermore, dynamic interactions between neural transcription factors appear in concert with a cell-type specific establishment of enhancer-promoter contacts (Bonev et al., 2017). Interestingly, also LADs undergo large repositioning across the genome (Peric-Hupkes et al., 2010). While increase in NL-contacts was generally associated with gene silencing, gene re-positioning to the nuclear interior correlated with their activation (Peric-Hupkes et al., 2010; van Steensel and Belmont, 2017). Globally, around half of all LADs were found to vary in different cell types leading to the discrimination between constant LADs (cLADs) and facultative LADs (fLADs) (Meuleman et al., 2013; Peric-Hupkes et al., 2010).

1.4 The nucleolus

The nucleolus is the largest nuclear sub-compartment of mammalian cells in which ribosomal RNA (rRNA) transcription, rRNA maturation and ribosome biogenesis take place. The membrane-less nucleolus is formed by clustered rRNA gene repeats located at nucleolar organizer regions (NORs), which are positioned at centromeric regions of five chromosomes in mouse cells (Dundr et al., 2000). The nucleolus can be further divided into several structural and functional subdomains. Transcription of rRNA genes catalyzed by RNA polymerase I (Pol I) occurs in fibrillar centers (FCs), while the subsequent modification and processing of the nascent pre-rRNA, as well as

its assembly in pre-ribosomal particles takes place in dense fibrillar components (DFCs). The third nucleolar domain, called granular component (GC), accounts for later rRNA processing steps and maturation of ribosomes (Scheer and Hock, 1999). Importantly, nucleolar structures are dynamically regulated in space and time. During mouse development, nucleolus formation occurs at the late 2-cell stage correlating with activation of Pol I transcription (Engel et al., 1977). In general, it is believed that formation and maintenance of nucleolar integrity is dependent on active transcription or rRNA genes. Accordingly, inhibition of Pol I transcription leads to rearrangement of nulceolar components followed by complete nucleolar disintegration (Nemeth and Grummt, 2018). Globally, nucleoli are in terms of transcription the most active domains of the mammalian genome, accounting for 35% of total transcription in proliferating cells (Moss et al., 2007). Thus, it is relatively surprising that the active nucleolus also serves as a hub for local heterochromatin localization (Nemeth and Langst, 2011). Loss of silencing correlates with rDNA instability, nucleolar disintegration and cellular senescence, suggesting essential roles of heterochromatin in nucleolar function and stability (Nemeth and Grummt, 2018). Conversely, the nucleolus participates in global heterochromatin organization within the nucleus. In differentiating cells, the it plays a pioneering role in shaping global heterochromatin formation and spreading (Savic et al., 2014). Furthermore, it was shown to be important for anchoring the inactive X chromosome to perinuclear compartments in female cells (Yang et al., 2015). In summary, the nucleolus is a transcriptionally highly active sub-compartment of the nucleus, which mainly accounts for ribosome biogenesis but also plays important roles in functional organization of the genome.

1.4.1 Ribosomal RNA (rRNA) genes

rRNA genes are organized in tandem repeat arrays, which – in mouse – are located on chromosomes 12, 15, 16, 18 and 19 (Dev et al., 1977). Each of these arrays can act as an independent NOR (Bell et al., 1992). In total, mice contain approximately 400 rRNA gene copies, each of around 45kb in size (Grozdanov et al., 2003). A single gene unit consists of a 13-14kb sequence encoding the pre-rRNA, which is separated by the long intergenic spacer (IGS) element of around 30kb. Regulatory elements, such as the core promoter element (CPE), upstream control element (UCE), spacer

promoter, enhancer repeats and the transcription terminator elements are located in the IGS (**Figure 12**) (Grob et al., 2011).

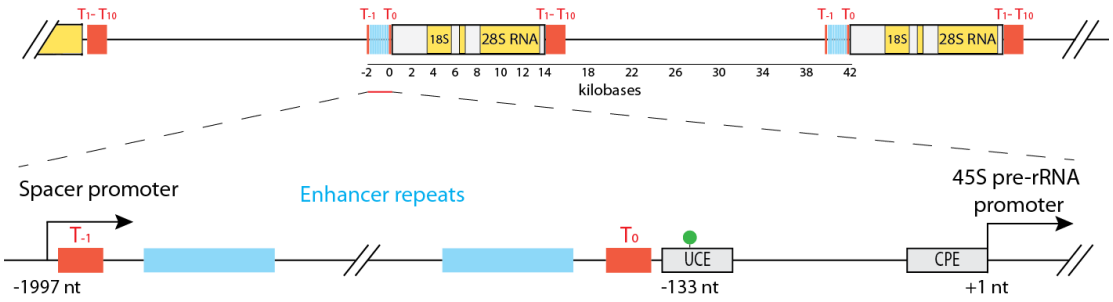


Figure 12. Structure of rRNA genes. RNA Pol I-mediated transcription is initiated at the spacer promoter giving rise to IGS-rRNA and at the core promoter driven by the core promoter element (CPE) for transcription of 45S pre-rRNA (indicated by black arrows). Red boxes mark the twelve terminator elements (T_{-1} to T_{10}). Cyan boxes indicate varying numbers of repetitive enhancer repeats. The upstream control element (UCE) contains a CpG at position -133, which is critical for rDNA silencing.

The mouse rRNA gene promoter consists of the CPE directly adjacent to the TSS and the UCE located 100bp upstream. The Upstream binding factor (UBF) specifically binds the CPE and UCE and is an important factor in promoting rDNA transcription. Together with the TATA binding protein (TBP) centered TIF-IB complex, UBF forms the pre-initiation complex required for recruitment of RNA Pol I (Moss et al., 2007). Strikingly, this complex remains associated to NORs during mitosis, bookmarking these loci for immediate re-activation in G1 (Leung et al., 2004). Active rDNA transcription generates the 45S pre-rRNA, which is further modified and processed into 28S, 18S and 5.8S rRNAs. These mature rRNA transcripts are subsequently packaged with ribosomal proteins forming the large and small subunits of ribosomes (Santoro, 2005). 2kb upstream of the core promoter resides the spacer promoter (Kuhn and Grummt, 1987). This alternative rDNA promoter displays several sequence homologies with the core promoter and is therefore also highly enriched for RNA Pol I occupancy (Santoro et al., 2010). Transcription from the spacer promoter gives rise to the 2kb lncRNA IGS-rRNA. In somatic cells IGS-rRNA is rapidly processed giving rise to a more stable 200nt lncRNA called promoter RNA (pRNA), which corresponds to the sequence to the core promoter element and is involved in rRNA gene silencing (Mayer et al., 2006; Santoro et al., 2010). The rDNA enhancer repeats are located between both the core and upstream promoter and are therefore also part of the IGS-rRNA. Various numbers in enhancer repeats (6, 9, 10, 11, 12 and 22) have been reported, of which a subclass containing 9 repeats is

particularly important for regulating the epigenetic state of rDNA (Santoro et al., 2010). Twelve copies of terminator elements (T_{-1}-T_{10}) represent additional *cis* regulatory sequences in the rDNA locus. T_0-T_{10} are located downstream of rRNA genes and are implicated in termination of rRNA transcription (Grummt et al., 1985). Two additional terminator elements are found upstream of the core promoter (T_0) and spacer promoter (T_{-1}), respectively. All Ts are specifically recognized and bound by transcription termination factor 1 (TTF1), which is implicated in promoting rDNA transcription and its termination (McStay and Grummt, 2008). However, bound to T_0 at the core promoter, TTF1 can also mediate to the recruitment of a repressor complex leading to the permanent silencing of the respective rRNA gene unit (Nemeth et al., 2004).

1.4.2 Chromatin and epigenetic regulation of rRNA genes

Despite the large demand for rRNA in mammalian cells, not all rRNA gene copies are actively transcribed. Visualization of rDNA genes by Miller spreads in *S. cerevisiae* revealed two types of rRNA genes: One class was actively transcribed by RNA Pol I, giving rise to the typical “Christmas tree” structures, while the other subclass did not show any indication of RNA synthesis (Miller and Beatty, 1969). Although Miller spreading techniques cannot be used for mammalian cells, the co-existence of active and silent rRNA genes within the same cell has been confirmed in higher eukaryotes (Santoro, 2011). Epigenetic characterizations revealed that genes harboring processive rRNA transcription are enriched for active histone marks such as H3K4me2 and H4ac (Santoro and Grummt, 2005; Santoro et al., 2002). Silent rRNA genes by contrast, are associated with heterochromatin protein 1 (HP1) and the silent marks H3K9me2, H3K27me3 and H4K20me3 (Santoro and Grummt, 2001, 2005; Santoro et al., 2002). Interestingly, most of these histone modifications were shown to peak at the spacer promoter without any detectable enrichment at the main promoter (Zentner et al., 2014). However, these findings must be treated with caution due to the low mappability of the rDNA locus (Zentner et al., 2014). In addition to the repressive histone marks, silent rDNA loci are also highly enriched for DNA methylation. Particularly one CpG at position -133 located within the UCE is directly involved in silencing rRNA transcription. When methylated, this CpG prevents the binding of UBF to rDNA chromatin and thereby impairs formation of the initiation complex

leading to transactional repression (Santoro and Grummt, 2001). Another important contribution to rRNA gene transcription comes from ATP-dependent chromatin remodeling complexes that regulate the exact nucleosome positioning. In mouse cells, a particular nucleosome at the core promoter element exists in two distinct positions, directly corresponding to the transcriptional state of the respective locus (Li et al., 2006). In active rRNA genes, this nucleosome occupies the sequence region from -157 to the TSS. In silent genes however, it extends over the TSS covering the region from -132 to +22. Situated in this "off" position, the -133 CpG becomes exposed at the 5' boundary of the nucleosome and can be targeted by DNMTs for DNA methylation, thereby stabilizing the silent state (Santoro and Grummt, 2001; Santoro et al., 2002). Specific recruitment of nucleosome remodeling complexes to the respective gene loci is mediated by TTF1 bound to the T_0 element (Langst et al., 1998; Langst et al., 1997). Importantly, the chromatin state of active and silent rRNA genes is stably propagated throughout the cell cycle ensuring epigenetic inheritance. These regulatory mechanisms controlling rDNA expression are not rRNA gene-specific but seem to act on a larger scale affecting entire rRNA gene clusters. Thus, active and silent NORs can be distinguished in higher eukaryotes (Pikaard, 2000; Schlesinger et al., 2009). While the exact mechanism mediating inheritance of active rRNA genes is not fully understood, many studies have highlighted the crucial role of the nucleolar remodeling complex (NoRC) in silencing rDNA chromatin in mammals (Santoro et al., 2002).

1.4.3 Nucleolar remodeling complex NoRC

NoRC was originally discovered by a yeast two-hybrid system screening for TTF1-interacting proteins. These analyses identified TTF1-interacting protein 5 (TIP5), which – in association with SNF2H – forms the ATP-dependent chromatin remodeling complex NoRC and binds to rRNA genes (Strohner et al., 2001). Further investigations revealed that NoRC is the key player in establishing and maintaining the heterochromatic state of silent rRNA genes (Li et al., 2006; Santoro et al., 2002; Strohner et al., 2001; Zhou et al., 2002). SNF2H is a member of the ISWI subfamily and serves as the catalytic subunit for several other chromatin remodeling complexes as well (Hota and Bruneau, 2016; Strohner et al., 2001). TIP5 on the other hand is a large, > 200kDa protein that is not known to occur in any other complex than NoRC

(Bochar et al., 2000; Bozhenok et al., 2002; Ito et al., 1999; LeRoy et al., 2000). TIP5 belongs to the protein family of bromodomain adjacent zinc finger (BAZ) proteins and has therefore also been alternatively named as Baz2a (Jones et al., 2000). While SNF2H is responsible the nucleosome remodeling activity, TIP5 provides the specificity for rDNA binding and serves as a binding platform for further factors required for transcriptional silencing. Thus, several TIP5 domains have been assigned crucial roles for mediating these functions. For instance, the C-terminal bromodomain allows binding of acetylated histones – H4K16ac in the case of rDNA – and is necessary to mediate HDAC1 recruitment. The PHD domain is required for the interaction with SNF2H, histone methyltrasnferases (HATs) and DNMTs (Zhou and Grummt, 2005; Zhou et al., 2002). The TIP5/ARBD/MBD (TAM) domain of TIP5 is of particular importance, since it binds stem loop structures of the lncRNA pRNA mediating itns rDNA-specific recruitment (Mayer et al., 2006). Immediately after replication of the silent rDNA in late S-phase, TIP5 is recruited to the rRNA promoter via its interaction with pRNA and T_0-bound TTF1 (Mayer et al., 2008; Savic et al., 2014). Through the recruitment of further factors, such as HDAC1, SETDB1, PARP1 and DNMTs, TIP5 represses rRNA transcription maintaining the silent state of rDNA chromatin (**Figure 13**) (Guetg et al., 2012; Li et al., 2006; Santoro and Grummt, 2005; Santoro et al., 2002; Zhou et al., 2002; Zhou et al., 2009).

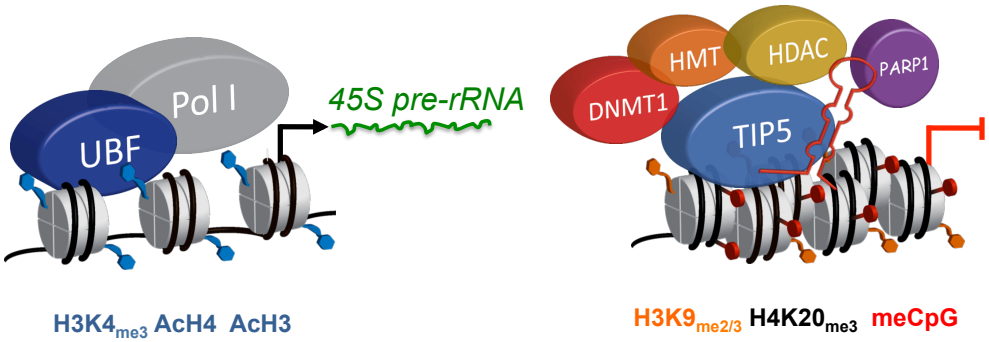


Figure 13. Schematic representation of active and silent rRNA genes. Transcriptionally active rRNA genes are occupied by UBF and RNA Pol I and are asociated by active histone marks (labeled in blue). Characterized by an open chromatin conformation, these rDNA repeats give rise to the 45S pre-rRNA. Chromatin of silent rDNA by contrast, is highly compact and marked by DNA methylation (red label) and silent histone PTMs (orange label). The silencing complex NoRC is recruited via the lncRNA pRNA associated to TIP5. Subsequent recruitment of histone methyltransferases (HMTs), HDACs and DNMTs mediate the rDNA silencing.

In addition to the recruitment of effector proteins, NoRC is also responsible for sliding of the rDNA promoter nucleosome into the "off" position covering the region -132 to +22 (Santoro, 2011). The following CpG methylation at position -133 and consequential displacement of UBF is required for efficient rRNA gene silencing, since its repression is blocked by treatment with 5-azacytidine (Santoro and Grummt, 2001). Taken together, NoRC has been identified as the main component responsible for rRNA gene silencing. Mechanistically, NoRC coordinates several enzymatic processes including ATP-dependent nucleosome remodeling, histone deacetylation, DNA and histone methylation to establish a closed chromatin structure and block rRNA transcription.

1.4.4 Long non-coding RNA pRNA

RNA Pol I-mediated transcription of the upstream promoter gives rise to the IGS-rRNA, which maps to the 2kb spacer region between both promoters and further contains sequences from the main promoter. IGS-rRNA transcription is highly cell cycle-regulated occurring only during early S-phase from a subset of active, hypomethylated rRNA genes that harbor nine enhancer repeats (Santoro et al., 2010). In mid-late S-phase, IGS-rRNA is processed to a shorter version that matches the rDNA promoter sequence and is therefore called promoter RNA (pRNA) (Santoro et al., 2010). The exact mechanism of IGS-rRNA processing is not fully understood but was shown to require the RNA helicase DHX9 (Leone et al., 2017). Interestingly, production of the ~200nt lncRNA pRNA coincides with NoRC recruitment to newly replicated, silent rRNA genes (Guetg et al., 2012; Li et al., 2006). A central pRNA sequence motif folds into a conserved stem loop structure that is specifically recognized by the TIP5-TAM domain. Binding to pRNA induces a conformational change of TIP5, which facilitates its interaction with other factors involved in rDNA silencing (Mayer et al., 2008). In support with this hypothesis pRNA is required for the interaction of TIP5 with TTF1 and PARP1, whereas its depletion via LNA causes defects in heterochromatin formation at rRNA genes (Guetg et al., 2012; Mayer et al., 2006; Savic et al., 2014). Further work elucidating the mechanism of NoRC recruitment revealed that single mutations in the pRNA loop structure but not in its 5' regions abolished the interaction of TIP5 with TTF1 and its association with rDNA (Mayer et al., 2008; Savic et al., 2014). Conversely, TIP5 TAM domain-mutants that

are impaired in pRNA binding fail to associate with chromatin and do not establish rDNA heterochromatin (Mayer et al., 2006). In summary, these studies suggest a mechanism by which rRNA genes are silenced via lncRNA-mediated recruitment of the chromatin remodeling complex NoRC.

Recent studies highlighted that pRNA-guided rDNA silencing does not only occur in a spatiotemporal-regulated manner throughout the cell cycle but is also tightly controlled during development. In line with their global lack in constitutive heterochromatin, compact chromatin structures at the nucleolus are absent in ESCs. Accordingly, all rRNA genes are active and devoid of heterochromatic marks (Savic et al., 2014). As ESCs exit pluripotency, TIP5 rapidly localizes to nucleoli and progressively establishes heterochromatin at rRNA genes during early differentiation processes (Savic et al., 2014). Strikingly, the impairment in rDNA silencing in ESCs cannot be explained by diminished NoRC expression, since TIP5 is even overexpressed in ESCs. By contrast, an absence of pRNA was shown to be the driving force for maintaining euchromatic rRNA genes in ESCs. Although ESCs actively express IGS-rRNA, it is not further processed into its mature form, resulting in a defect in NoRC recruitment to rDNA. In agreement with this observation, transfection of mature pRNA is sufficient for directing TIP5 to rDNA followed by heterochromatin formation at these loci in ESCs (Savic et al., 2014). Moreover, it was shown that TIP5 is able to bind IGS-rRNA but thereby inhibits its interaction with TTF1. Accordingly, immunofluorescence experiments in ESCs revealed that TIP5 is even excluded from nucleoli under basal conditions, whereas transfection of pRNA rapidly leads to its exclusive nucleolar localization (Leone et al., 2017; Savic et al., 2014). Strikingly, pRNA administration causes chromatin compaction not only at the rDNA but leads to global heterochromatin formation also at non-nucleolar regions, such as major and minor satellite repeats. This re-organization of the genome is accompanied by a global increase in the repressive H3K9me2 levels and by a loss in the pluripotent potential (Savić et al., 2014). Thus, the nucleolus functions – in addition to its known role in ribosome production – as a seeding point for global heterochromatin formation leading to re-structuring of the genome in differentiating cells. Furthermore, these observations imply that NoRC undergoes a crucial switch-of-function from ESCs to differentiated cells. While its role in silencing rRNA genes in differentiating and somatic cells has been extensively investigated, the high expression of TIP5 in ESCs suggests a non-nucleolar function in pluripotency.

2. Aims

Recent results of our laboratory suggested an important role for the nucleolar remodeling complex (NoRC) in ESCs (Savic et al., 2014). NoRC consist of two subunits, the ATPase SNF2H/SMARCA5 and TIP5 and has previously been identified as the key determinant for the establishment of epigenetic silencing at ribosomal rRNA genes in somatic cells (Li et al., 2006; Santoro et al., 2002; Strohner et al., 2001; Zhou et al., 2002). In embryonic stem cells (ESCs) however, all rRNA genes are actively transcribed and therefore devoid of heterochromatic blocks (Savic et al., 2014). Nevertheless TIP5 – the main component of NoRC – is highly expressed in ESCs, but does not associate with rDNA and, hence, does not regulate rRNA gene transcription. Upon differentiation, TIP5 is rapidly recruited to rDNA by a lncRNA-mediated mechanism and progressively establishes permanent silencing of a subset of rRNA genes (Savic et al., 2014). Thus, these results suggest that NoRC rapidly switches its function during the transition of ESCs into lineage-committed cells. Remarkably, ESCs depleted of TIP5 grow slower and have an impaired ability to differentiate, highlighting an essential role of NoRC in pluripotency that is not related to the regulation of rRNA genes (Savic et al., 2014). The aim of this project was to mechanistically elucidate the role of NoRC in ESCs. Given the two ESCs models in serum and 2i, we further aimed to determine whether NoRC undergoes an additional switch-of-function from ground state to primed pluripotency. Understanding how the pluripotent state of ESCs is maintained has the potential to help future clinical applications of regenerative medicine. Moreover, given the recognized similarity of ESCs and cancer cells, increased knowledge of ESC biology might also lead to new approaches for cancer therapies. The recent link established in our lab concerning the role of TIP5 in ESCs and prostate cancer cells has the potential to directly improve therapeutic and prognostic strategies in cancer and regenerative medicine.

3. Results

3.1 Research articles

3.1.1 NoRC complex is a regulator of chromatin architecture of ground-state pluripotent cells

Authors: **Damian Dalcher**, Eva Vollenweider, Marc W. Schmid, Jennifer Yihong Tan, Valerio Bianchi, Wouter de Laat & Raffaella Santoro

Journal: Manuscript under reviewing in Cell Stem Cell

Contribution: D.D. contributed to all figures and wrote the manuscript with RS.

3.1.2 Establishment of heterochromatin at rRNA genes is required for embryonic stem cell differentiation

Authors: Sergio Leone, Dominik Bär, Coenraad Frederik Slabber, **Damian Dalcher** & Raffaella Santoro

Journal: EMBO Reports

DOI: 10.15252/embr.201744330

External link: http://embor.embopress.org/content/18/7/1248.long

Contribution: D.D. contributed to Figure EV1

3.1.3 Pramel7 mediates ground-state pluripotency through proteasomal–epigenetic combined pathways

Authors: Urs Graf, Elisa A. Casanova, Sarah Wyck, **Damian Dalcher**, Marco Gatti, Eva Vollenweider, Michal J. Okoniewski, Fabienne A.Weber, Sameera S. Patel, MarcW. Schmid, Jiwen Li, Jafar Sharif, Guido A.Wanner, Haruhiko Koseki, JieminWong, Pawel Pelczar, Lorenza Penengo, Raffaella Santoro and Paolo Cinelli

Journal: Nature Cell Biology

DOI: 10.1038/ncb3554.

External link: https://www.nature.com/articles/ncb3554

Contribution: D.D. contributed to figures 1, 2 and S4

NoRC complex regulates chromatin architecture of ground state pluripotent stem cells

Damian Dalcher[1,2], Eva Vollenweider[1,2], Jennifer Yihong Tan[3], Marc W. Schmid[4], Valerio Bianchi[5], Wouter de Laat [5] and Raffaella Santoro[1#]

1. Department of Molecular Mechanisms of Disease, DMMD, University of Zurich, 8057 Zurich, Switzerland

2. Molecular Life Science Program, Life Science Zurich Graduate School, University of Zurich, 8057 Zurich, Switzerland

3. Department of Computational Biology, University of Lausanne, 1015 Lausanne, Switzerland

4. Service and Support for Science IT, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

5. Hubrecht Institute-KNAW and University Medical Center Utrecht, 3584 CT Utrecht, the Netherlands

Corresponding author:

Raffaella Santoro (raffaella.santoro@dmmd.uzh.ch)

Abstract

3D genome architecture is important for gene regulatory circuits. Yet, its contribution to early developmental states remains unclear. Here we show that ground state and developmentally primed pluripotent embryonic stem cells (ESCs) differ in their dependency on a set of chromatin regulators. One of these factors is the chromatin remodeling complex NoRC, composed of TIP5 and SNF2H. Although NoRC is highly expressed in both pluripotency states, only ground state ESCs depend on NoRC for correct gene expression, control of global H3K27me3 occupancy, growth and differentiation. NoRC binds large and active regions within the A-compartment, interacts with topoisomerase 2A (TOP2A) and cohesin, and limits far-*cis* chromatin contacts in order to maintain ground state 3D genome architecture. NoRC acts together with TOP2A as depletion or activity inhibition of TOP2A phenocopies the changes in gene expression and H3K27me3 occupancy observed upon TIP5 knockdown in ground state ESCs without any influence in ESC+serum. These findings highlighted different degrees of genome dynamics between the early phases of development through the action of chromatin remodelers and the relief of chromatin torsional stress.

Introduction

Embryonic stem cells (ESCs) are derived from pre-implantation epiblasts of the late blastocyst and have the potential to produce all tissues when injected into host embryos (Boroviak et al., 2014; Nichols and Smith, 2009). These properties make ESCs an invaluable system to elucidate regulatory mechanisms during early mammalian development. Compared to differentiated cells, ESCs display a less condensed and more euchromatic conformation. This structure is particularly malleable and transcriptionally permissive and reflects the plastic state of pluripotent cells that must retain the ability to enter into any lineage specification (de Wit et al., 2013; Gaspar-Maia et al., 2011b; Gorkin et al., 2014). The positive correlation between epigenetic repressed structures and more advanced developmental stages is also apparent between ground state and developmentally primed pluripotent ESCs (Ficz et al., 2013; Habibi et al., 2013; Marks et al., 2012). ESCs exist in a variety of states *in vitro*, largely defined by the culture conditions, which reflect the natural development of the embryo from the blastocyst to post-implantation stages (Hackett and Surani, 2014). ESCs can be

propagated in medium containing fetal calf serum and leukaemia inhibitory factor (LIF) (ESC+serum) or in serum-free 2i medium (ESC+2i) that contains LIF plus two small-molecule kinase inhibitors for MEK/ERK (PD0325901) and GSK3 (CHIR99021) (Ying et al., 2008). Both ESC+2i and ESC+serum are pluripotent. However, ESC+2i closely resemble the developmental ground state *in vivo* whereas ESC+serum are developmentally primed (Boroviak et al., 2014). Accordingly, ESC+2i have a higher efficiency in chimera formation compared to ESC+serum (Alexandrova et al., 2016). Moreover, ESC+2i and ESC+serum display distinct transcriptional profiles, epigenetic landscapes and chromatin structures. In comparison to ESC+serum, ESC+2i show decreased expression of lineage priming genes (Marks et al., 2012). Furthermore, the epigenetic and chromatin organization of ESC+2i is in a less repressed state. ESC+2i have hypomethylated DNA similar to pre-implantation embryos, whereas ESC+serum genome is hypermethylated, reminiscent of post-implantation embryos (Ficz et al., 2013; Habibi et al., 2013; Leitch et al., 2013; Marks et al., 2012). Similarly, there is a reduced prevalence of the repressive H3K27me3 mark at Polycomb target promoters in ESC+2i (Joshi et al., 2015; Marks et al., 2012). Finally, quantitative super-resolution nanoscopy coupled with computer simulations revealed that ESC+2i harbor a more open chromatin configuration when compared to ESC+serum (Ricci et al., 2015).

The 3D genome organization has been thought to be functionally important for correct execution of gene expression programs. The crucial roles in chromatin organization of architectural proteins, notably cohesin and the transcriptional insulator CTCF, have been extensively studied (Dixon et al., 2012; Guo et al., 2015; Merkenschlager and Nora, 2016; Rao et al., 2017; Schwarzer et al., 2017). Cohesin and CTCF co-localize at topology-associated domain (TAD) boundaries (Dixon et al., 2012; Rao et al., 2014). Data supported a role of CTCF as boundary element whereas cohesin is thought to be loaded at sites of active transcription and travels along the chromosome arms by a loop extrusion mechanism until reaching convergently oriented CTCF sites (de Wit et al., 2015; Guo et al., 2015; Nora et al., 2017; Sanborn et al., 2015). Recent studies have also suggested crucial roles of Polycomb repressive complexes 1 and 2 (PRC1, PRC2) in the establishment of long-range intra- and interchromosomal interactions in ESCs (Joshi et al., 2015; Schoenfelder et al., 2015b). Importantly, these long-range interactions were detected only in ESC+serum and not in

ESC+2i, suggesting that PRC-mediated epigenetic repressive signature is a determinant of 3D genome organization during early development. However, how the 3D organization of chromatin is regulated and how it contributes to ground state and developmental primed ESCs remains unclear.

Chromatin remodeling complexes use the energy of ATP hydrolysis to move, destabilize, eject, or restructure nucleosomes (Clapier and Cairns, 2009). Although their role on local chromatin changes has been intensively investigated, the possibility whether they can shape 3D genome architecture remains to be elucidated. The nucleolar remodelling complex NoRC is a chromatin remodeling complexes, which consists of two subunits, the ATPase SMARCA5/SNF2H and TIP5, a >200 kDa protein that shares sequence homology with the largest subunits of SNF2H/ISWI-containing remodeling complexes (Ralf Strohner, 2001; Santoro et al., 2002). In differentiated cells, TIP5 is mainly localized in nucleoli, associates with ribosomal RNA (rRNA) genes and establishes their epigenetic silencing (Guetg et al., 2010; Mayer et al., 2006; Santoro et al., 2002). In contrast, in ESCs the ability of NoRC to silence rRNA gene transcription is prevented through long non-coding RNA (lncRNA)-mediated mechanisms that result in the impairment of TIP5 recruitment to rRNA genes (Leone et al., 2017; Savić et al., 2014). As a consequence, all rRNA genes in ESCs are active and display euchromatic features, including the lack of DNA methylation (Savić et al., 2014; Schlesinger et al., 2009).

The ability of NoRC to act as a repressor of rRNA gene transcription only upon exit of pluripotency suggests a switch of function during development. In this work, we explored the role of NoRC in ground state and primed pluripotency states and found an unexpected non-nucleolar function of NoRC in ESCs. We showed that NoRC is a regulator of chromatin architecture of ground state pluripotent ESCs and contributes to local chromatin compaction by limiting far-*cis* contacts. NoRC associates with large and active regions of the genome and interacts with DNA topoisomerase 2A (TOP2A) and cohesin. Our results further indicate that NoRC is required to regulate the expression of developmental genes and to control global H3K27me3 occupancies. Furthermore, the data underscored fundamental differences in the chromatin organization between ground state and primed ESCs. While ESC+2i depend on NoRC for maintenance of correct gene expression and H3K27me3 signature, proliferation

and differentiation capacities, ESC+serum do not. The results indicated that the open structure of ground state chromatin strongly depends on chromatin remodelers and on the relief of chromatin torsional stress in order to reinforce active and repressed genome partitioning through chromatin compaction. In contrast, the more repressed and closed chromatin of ESC+serum is less dependent on factors that additionally limit chromatin dynamics and flexibility. These findings highlight how such closely and developmentally related ESC types display different requirements for the organization of genome architecture, shedding light on crucial chromatin dynamics taking place during early mammalian development.

Results

Ground state and primed pluripotent ESCs differ in the requirement of NoRC activity for cell proliferation and differentiation capacity

Our initial analysis suggested that in ESCs NoRC function was not related to rRNA transcriptional control and was distinct between between ground state and primed ESCs. First, TIP5 - the main NoRC component - is more highly expressed in ESC+2i than in differentiated cells (**Fig. 1A**). Second, although not bound to rRNA genes, TIP5 is still tightly associated with chromatin in ESC+2i (**Fig. 1B**). Third, proliferation of ESC+2i decreased upon TIP5 depletion by siRNA (**Fig. 1D,E**). Similar results were obtained with a different siRNA and in another ESC line (**Supplementary Figure 1A,B**). Fourth, and consistent with previous results (Savić et al., 2014), after induction of monolayer differentiation upon withdrawal of LIF, ESC+2i treated with siRNA-*Tip5* underwent cell death while control cells displayed morphological structures typical of differentiated cells and were negative for alkaline phosphatase staining (**Fig. 1F,G**). All these results indicated that ground state ESCs depend on TIP5 expression for proliferation and differentiation capacity and highlighted an unexpected non-nucleolar function of NoRC. Next we asked whether proliferation and differentiation capacities of primed ESC+serum were also dependent on NoRC. Surprisingly, although TIP5 expression levels and knockdown efficiency in ESC+2i and ESC+serum were similar (**Fig. 1C,D**), depletion of TIP5 in ESC+serum did not cause any evident defect in proliferation or differentiation (**Fig. 1E-G**). Taken together these results highlighted a

substantial difference in the requirement of NoRC for cell proliferation and differentiation capacities between ground state and primed pluripotent ESCs.

In ESC+2i NoRC regulates genes with a bivalent chromatin signature

Previous data demonstrated that ground state and primed pluripotent ESCs display significant differences in their transcriptional profiles and epigenetic landscapes (Ficz et al., 2013; Habibi et al., 2013; Marks et al., 2012). The unexpected finding that only ESC+2i rely on NoRC for cell proliferation and their differentiation capacity suggests that NoRC contributes to the molecular differences between these two pluripotency states. To understand these distinct phenotypes in more detail we performed transcriptome analyses of ESC+2i and ESC+serum treated with siRNA-*Tip5* or siRNA-control (**Fig. 2A, Supplementary Table 1**). We found that depletion of TIP5 induces higher differential gene expression in ESC+2i than in ESC+serum. Upon TIP5 depletion in ESC+2i, 1934 genes showed transcriptional changes (log_2 fold change = 0.58; P<0.05; 1236 upregulated and 698 downregulated). In contrast, in ESC+serum the number of genes affected by TIP5 knockdown was lower (351 upregulated and 207 downregulated) and only a minority of them were upregulated (65) or downregulated (56) by TIP5 knockdown in both ESC states (**Supplementary Figure 2A**). More than half of genes (1052) regulated by TIP5 in ESC+2i were also differentially expressed between ESC+2i and ESC+serum (**Supplementary Figure 2B**). The top 10 gene ontology (GO) terms for genes regulated by TIP5 in ESC+2i revealed a high enrichment in pathways linked to developmental processes (**Fig. 2B, Supplementary Table 2**). This was not the case for genes regulated by TIP5 in ESC+serum, which are mostly implicated in biological processes linked to cell signalling and cell adhesion. To determine which epigenetic signature characterizes TIP5-regulated genes in ESC+2i, we performed ChIPseq analysis in ESC+2i for the repressive H3K27me3 and the active H3K4me3 histone modifications (**Fig. 2C,D**). Although H3K27me3 levels at gene promoters in ESC+2i are generally lower than in ESC+serum (**Fig. 2D**) (Marks et al., 2012), we found that promoters of TIP5-regulated genes were enriched for the repressive histone mark H3K27me3 and contained the active H3K4me3 modification. This epigenetic signature characterizes bivalent domains, which are considered to poise expression of developmental genes, allowing timely activation while maintaining

repression in the absence of differentiation signals (Voigt et al., 2013). Compared to random genes, H3K27me3 levels at TIP5-regulated genes transcription start site (TSS) in ESC+2i were higher whereas H3K4me3 were lower (**Fig. 2C**). Moreover, in ESC+2i these promoters had high enrichment of Ezh2, Suz12 and Ring1B, components of Polycomb complexes 1 and 2 (PRC1, PRC2) (**Supplementary Figure 2C**). Expression levels of TIP5-regulated genes were low in both ESC+2i and ESC+serum (**Fig. 2E**), a result that is consistent with the lower occupancy of the active mark H3K36me3 across their gene body and RNA Polymerase II at their TSS compared to random genes (**Fig. 2C, Supplementary Figure 2C**). Validation by qRT-PCR supported a role of TIP5 in the regulation of these genes in ESC+2i but not in ESC+serum without affecting pluripotency genes *Rex1* and *Nanog* (**Fig. 2F**). Similar results were obtained using a different siRNA-*Tip5* sequence and another ESC line (**Supplementary Figure 2D,E**). Taken together, these results indicate that only ground state ESCs depend on TIP5, the component of NoRC, to regulate expression of bivalent-marked genes implicated in developmental processes. By contrast, developmentally primed ESCs seemed to be unaffected by changes in TIP5 expression levels.

NoRC regulates H3K27me3 occupancy in ground state ESC

Previous studies have shown that ESC+serum are epigenetically more restricted and developmentally primed compared to ground state ESC+2i, which in turn are characterized by reduced genome wide DNA methylation and H3K27me3 at Polycomb target promoters (Ficz et al., 2013; Habibi et al., 2013; Leitch et al., 2013; Marks et al., 2012). To determine whether NoRC has a specific function in the regulation of ground state epigenetic signature, we analysed occupancy of active histone marks H3K4me3 and H3K27ac and the repressive H3K27me3 modification upon TIP5 knockdown. Western blot analyses indicated that total levels of these histone modifications were not altered in the absence of TIP5 (**Fig. 3A, Supplementary Figure 3A**). Surprisingly, ChIPseq analysis revealed a global re-distribution of H3K27me3 in ESC+2i but not in ESC+serum upon TIP5 depletion (**Fig. 3B, Fig. 4B, Supplementary Figure 3B**). In contrast, no changes were detected for H3K4me3 and H3K27ac at TSSs upon TIP5 knockdown (**Fig. 3B, Supplementary Figure 3B**). For genes with high levels of H3K27me3 (first quartile, Q1) in ESC+2i, TIP5 depletion resulted in a

decrease of H3K27me3, while genes with relatively low H3K27me3 levels (quartiles 2-4, Q2-4) displayed an increase of this repressive mark. ChIP-qPCR experiments validated the decrease of H3K27me3 at TSS of selected genes with high H3K27me3 content and showed that this effect was specific for ESC+2i (**Fig. 3C**). A similar result was obtained with a different siRNA-*Tip5* sequence (**Supplementary Figure 3C**). Importantly, redistribution of H3K27me3 was not only restricted to TSS but was also detected at a subset of regions in ESC+2i where CTCF sites mark transitions in H3K27me3 enrichment (**Fig. 3D, E).** Upon TIP5 depletion, H3K27me3 redistributed over CTCF sites, decreasing at high H3K27me3 regions and increasing at adjacent low H3K27me3 regions. In contrast, H3K27me3 transitions marked by CTCF sites in ESC+serum were not affected upon TIP5 knockdown. We validated these results with independent H3K27me3 ChIP-qPCR experiments and another siRNA-*Tip5* (**Fig. 3E, Supplementary Figure 3B**). In summary, these results highlight a difference in chromatin organization between ground state and developmental primed ESCs, implying that only ESC+2i require NoRC to reinforce H3K27me3 boundaries.

To determine whether the specific requirement of NoRC in ground state is related to the hypomethylated genome and low H3K27me3 occupancy at PRC targets that characterizes ESC+2i, we analysed the role of TIP5 in controlling gene expression and H3K27me3 deposition in DNA methyltransferases triple knockout ESCs (DNMT1, 3B and 3A KO, TKO-ESC) and ESCs knockout for core components of PRC1 and PRC2 complexes (Ring1b$^{-/-}$ ESC and Eed$^{-/-}$ ESC). TIP5 knockdown in TKO-ESCs cultured in serum did not affect cell proliferation, expression of TIP5-regulated genes and H3K27me3 promoter occupancy, indicating that the specific function of TIP5 in ESC+2i does not depend on a low DNA methylation content (**Supplementary Figures 3D-F**). Analyses of Eed$^{-/-}$ and Ring1b$^{-/-}$ ESCs cultured in 2i showed that TIP5 depletion still affects cell proliferation and gene expression (**Supplementary Figures 3G-I**). These results are also in agreement with previous studies showing that global expression of bivalent genes is not altered in the absence of H3K27me3 in ESC+2i (Galonska et al., 2015a). Thus, although TIP5 is required for a correct distribution of H3K27me3 in ESC+2i, alterations in H3K27me3 alone are not sufficient to explain TIP5 impact on cell proliferation and transcriptional changes. Taken together these results suggest that alterations in gene expression and H3K27me3 are two independent readouts of NoRC

dependent changes of specific ground state chromatin conformation.

TIP5 associates with large and active genomic regions of ground state ESCs

The results described above suggested that the chromatin architecture of ground state ESCs requires NoRC to maintain correct gene expression and H3K27me3 levels. Since TIP5 is bound to chromatin (**Fig. 1B**), we analysed and compared TIP5 genomic occupancy in ESC+2i and ESC+serum. We established an ESC line containing a FLAG-HA (F/H) tag at the N-terminus of the endogenous TIP5 through CRISPR/Cas9 technology (**Fig. 4A**). Previous studies showed that the fusion of the F/H peptide at the N-terminus of TIP5 does not affect its activity (Guetg et al., 2012; Mayer et al., 2006). The obtained ESC lines (F/H-TIP5-ESC) contain F/H insertion at both *TIP5* alleles and express TIP5 at levels similar to wild-type (wt) cells (**Supplementary Fig. 4A-C**). We performed ChIPseq analysis with anti-FLAG or anti-HA immunoprecipitation and assessed the specificity of TIP5-ChIP by comparing FLAG- or HA-ChIP of F/H-TIP5-ESCs and wt-ESCs (**Fig. 4**). Surprisingly, the genomic binding pattern of TIP5 in both ESC+2i and ESC+serum did not appear as a distinct peak-like profile, but was rather enriched over large domains that extended up to several hundred kb (**Fig**. **4B**)**.** Although TIP5 genomic distribution was similar in ESC+serum and ESC+2i, TIP5 binding was higher in ESC+2i (**Fig. Fig. 4B-D; Supplementary Fig. 4D**). We validated the levels of TIP5 binding at selected genomic regions of ESC+2i and ESC+serum by classical ChIP-qPCR experiments (**Figure 4D**). The ChIP results were also consistent with previous data showing a lack of TIP5 association with rRNA genes in both ESC+2i and ESC+serum (Savić et al., 2014).

In contrast to differentiated cells where TIP5 associates with repressive epigenetic signatures (i.e. DNA methylation and K3K9me2/3) (Santoro et al., 2002), we found that TIP5 occupies active regions of the genome (marked by H3K4me1 and H3K27ac) and poorly correlates with H3K27me3 and H3K9me3 regions in ESC+2i (**Figure 4E**). Accordingly, TIP5-bound regions were found in internal chromatin domains (ICDs) and were excluded from the repressed lamina-associated domains (LADs) (Peric-Hupkes et al., 2010) (**Supplementary Fig. 4E**). Importantly, TIP5-regulated genes in ESC+2i did not show a preferential association with TIP5 (**Supplementary Fig. 4F**), an expected result given the high H3K27m3 levels at these

genes (**Fig. 2C**). Analysis at TSS of all TIP5-bound genes revealed that they were enriched for the active H3K4me3 and H3K27ac marks and depleted in H3K27me3 (**Fig. 4F**). Moreover, in ESC+2i but not in ESC+serum H3K27me3 at TIP5-bound genes – usually low at promoters of these genes –increased upon knockdown of TIP5 whereas H3K4me3 levels were not affected (**Fig. 4G**). These results suggest that NoRC does not directly regulate gene expression or H3K27me3 occupancy, since neither TIP5-regulated genes nor regions with high H3K27me3 were specifically bound by TIP5. Instead, NoRC appears to function as a structural factor presumably shaping active chromatin domains of ground state pluripotent cells.

TIP5 interacts with TOP2A and cohesin to regulate gene expression and H3K27me3 occupancy in ESC+2i

In order to understand how TIP5 shapes and specifically regulates chromatin of ESC+2i, we analysed which proteins interact with TIP5. Due to its tight association with chromatin, we aimed to purify TIP5 and its direct interaction partners in their native environment by establishing a protocol that allows immunoprecipitation from purified chromatin (Chromatin-IP, **Fig. 5A**). F/H-TIP5 and wt-ESC nuclei were incubated with the reversible protein-protein specific crosslinker dithiobis(succinimidyl propionate) (DSP) and chromatin was isolated by centrifugation of nuclear extracts. After digestion with MNase, solubilisation of mononucleosomes was achieved with 1% SDS, which does not affect protein-protein interactions stabilized by DSP crosslinking. The identification of TIP5-interacting proteins on chromatin was determined by comparing anti-FLAG immunoprecipitates from F/H-TIP5 and wt-ESC chromatin followed by mass spectrometry (**Fig. 5B, Supplementary Table 3**). Analysis from three independent experiments revealed 24 proteins consistently interacting with TIP5 in ESC+2i. As expected, the strongest interacting partner of TIP5 was SNF2h (**Fig. 5B**), indicating that the function of TIP5 in ESCs is directed via the NoRC complex. Moreover, we detected all known and previously validated TIP5-interactors such as DNMT1, PARP1 and DHX9 (Guetg et al., 2012; Leone et al., 2017; Ralf Strohner, 2001; Santoro et al., 2002). Interestingly, among the top TIP5-interacting proteins we found Topoisomerase 2A (TOP2A) and structural maintenance of chromosomes protein 3 (SMC3), which is part of the cohesin

complex (**Fig. 5B,C**). Furthermore, two out of three chromatin IP experiments detected the interaction of TIP5 with another cohesin component, SMC1 (**Supplementary Table 3**). Both TOP2A and cohesin are well known regulators of chromatin architecture. Topoisomerases exert the key function in relieving torsional stress of DNA (Nitiss, 2009). In ESCs, TOP2A is the most highly expressed type II isoenzyme that catalyzes the transient passage of two DNA duplexes and its inactivation is embryonically lethal (Akimitsu et al., 2003; Tiwari et al., 2012). Cohesin proteins are key architectural components of the genome that anchor long-range interactions allowing the formation of sub-megabase domains (TADs) (Schwarzer et al., 2017).

To determine whether TOP2A and cohesin mediate TIP5 function in ESC+2i, we analysed and compared gene expression in ESC+2i and ESC+serum upon knockdown of TOP2A and SMC3 or after treatment with ICRF-193, a potent TOP2 inhibitor (Pommier et al., 2010). Knockdown of TOP2A and SMC3 or catalytic inhibition of type 2 topoisomerases in ESC+2i induced transcriptional changes at candidate genes that are highly similar to those observed upon TIP5 knockdown (**Fig. 6A-C**). Moreover, the lack of apparent additive activation or repression of TIP5-regulated genes upon combined depletion suggests that TIP5 regulates gene expression in ESC+2i with TOP2A or SMC3 via shared pathways. Importantly, depletion of TOP2A or treatment with ICRF-193 in ESC+serum did not affect gene expression at any of the analyzed TIP5-regulated genes. SMC3 knockdown in ESC+serum, however, could still induce activation of few genes as found in ESC+2i (*Sema6, Lef1* and *Efnb2*). Therefore, as for NoRC, ESC+2i require TOP2A and SMC3 for correct gene expression, further highlighting substantial differences in the chromatin architecture of ground state and primed pluripotent cells.

To determine whether TOP2A or SCM3 regulate H3K27me3 occupancy, we performed ChIP analyses in ESC+2i and ESC+serum. Similarly to TIP5 knockdown, upon TOP2A depletion ESC+2i but not ESC+serum decreased H3K27me3 levels at both TIP5-regulated genes and regions where CTCF sites marked transitions in H3K27me3 enrichment (**Fig. 6D, E**). In contrast to TIP5 and TOP2A, SMC3 knockdown in ESC+2i did not cause evident changes in H3K27me3 occupancy at any of the analyzed regions (**Fig. 6F**). This result is consistent with a recent study showing that H3K27me3 is unaffected by cohesin (Rao et al., 2017). To

determine how TOP2A and SMC3 act together with NoRC, we analyzed whether they are required for the association of TIP5 with chromatin. We performed TIP5 ChIP analyses in ESC+2i and found that TOP2A or SMC3 depletion as well as inhibition of topoisomerase activity decreased the association of TIP5 with TIP5-bound sequences (**Fig. 6G-I**). Taken together these findings indicate that NoRC – together with TOP2A and cohesin –regulates gene expression specifically in ESC+2i. Furthermore, the data suggest that the release of chromatin torsional stress mediated by TOP2A is particularly required in ground state ESC to maintain correct H3K27me3 occupancy.

NoRC is required for 3D genome architecture of ground state ESCs

The results described so far point to a role of NoRC in chromatin organization of ground state ESCs. To determine whether NoRC affects chromatin architecture in ESC+2i, we generated DNA-DNA contact maps in control and TIP5 depleted cells using chromosome conformation capture (HiC) analysis (Bonev and Cavalli, 2016). We generated roughly 200 million valid read pairs from ESC+2i treated with siRNA-control (95 million) and siRNA-*Tip5* (105 million). As expected, TIP5 associates with the A compartment of ESC+2i (**Fig. 7A**). Globally, we did not detect any changes in the organization of the genome into A/B compartments upon TIP5 depletion (**Supplementary Fig. 5A**). However, we observed a genome wide increase in far-*cis* contacts (>10Mb) in ESCs lacking TIP5 (**Fig. 7B, Supplementary Fig. 5B**), suggesting a role of NoRC in regulating global genome architecture. These results indicate that NoRC limits long-range chromatin contacts by mediating local chromatin compaction, an action that recent work has suggested to be important to reinforce active and repressed genome partitioning (Haarhuis et al., 2017; Schwarzer et al., 2017). The defect in chromatin organization was also evident at the H3K27me3 domains *HoxA* and *HoxB* loci (**Fig. 7C-D**). In ESCs, the *HoxA* gene cluster shows strong internal interactions, forming a small subcompartment that separates two larger TADs (**Fig. 7C**, left panel). In the absence of TIP5 the contacts within this subTAD became weak and the entire cluster fused with two neighboring TAD structures. Concomitant with the loss of interactions within the *HoxA* locus, new contacts were also visible in the adjacent TADs. Interestingly, upon TIP5 knockdown in ESC+2i, *HoxA1* and *HoxA7* were upregulated whereas *Evx1* was downregulated (**Fig. 2F,**

Supplementary Table 1), supporting a structural role of NoRC in preventing the separation of *HoxA* locus into two compartments. Similar structural changes were also observed at the *HoxB* locus, with the loss of loop anchor points and fusion with the adjacent compartment (**Fig. 7D**, right panel). Further visual inspection of Hi-C maps revealed that ESCs depleted of TIP5 displayed several structural alterations, such as loss and gain of loop anchor contacts (**Supplementary Fig. 5C-D**). In summary, we propose an active role of NoRC in controlling global chromatin architecture, which is particularly important for maintaining correct gene expression signatures and epigenetic landscapes in ground state ESC+2i.

Discussion

In this study we show that ground state and primed ESCs display substantial differences in their chromatin organization due to their differential dependency on the chromatin remodeling complex NoRC. Upon TIP5 depletion - the main component of NoRC - ground state ESCs proliferated slower and displayed highly impaired differentiation. Furthermore, we demonstrated that NoRC is required for maintenance of correct gene expression signatures and H3K27me3 occupancy in ESC+2i. Importantly, this strong dependency on NoRC was ground state-specific, since TIP5 depletion in ESC+serum did not result in any apparent defects. Strikingly, we could recapitulate these ground state-specific phenotypes by depleting or inactivating the chromatin regulators TOP2A and SMC3, which we found associated with TIP5 on chromatin.

3D genome architecture has been proposed to be highly important for gene regulatory circuits. Recent work revealed that, relative to ESC+serum, ESC+2i have a less condensed chromatin structure and lack long-range genomic interactions (Joshi et al., 2015; Ricci et al., 2015). Moreover, in comparison to ESC+serum, ESC+2i contain more CTCF-occupied sites and loop structures (Beagan et al., 2017), suggesting that the 3D chromatin organization plays a particularly important role in ground state ESCs. Our results indicated that NoRC is involved in regulating genome architecture in ground state ESCs. We identified TOP2A and cohesin as strong TIP5 interacting partners on chromatin and showed alterations in far-*cis* genome contact frequencies upon TIP5 knockdown in ESC+2i. Interestingly, we did not detect CTCF as an interacting partner of TIP5, suggesting that cohesin might associate with

NoRC while traveling along the chromosome arms rather than at the stable anchor points co-occupied by CTCF. Accordingly, we did not detect any TIP5 enrichment at CTCF/cohesin peaks (data not shown). Our HiC analysis revealed global changes in far-*cis* contacts (>10Mb) but does not provide the required resolution for elucidating changes in local chromatin structures. However, a recent study highlighted that abolishing cohesin loading resulted in a highly similar increase in these far-*cis* chromatin contacts (Haarhuis et al., 2017). Since cohesin is known to limit local chromatin dynamics (Nozaki et al., 2017), it has been suggested that local chromatin compaction globally limits long-range chromatin contacts and therefore reinforces active and repressed genome partitioning (Haarhuis et al., 2017; Schwarzer et al., 2017). Strikingly, the reciprocal experiment of restricting cohesin removal from chromatin revealed the inverse outcome resulting in a reduction of far-*cis* chromatin contacts, further supporting this hypothesis (Haarhuis et al., 2017). Thus, our results suggest that NoRC contributes to local chromatin compaction and thereby limits far-*cis* contacts in ground state ESCs. Chromatin of ESC+serum by contrast, generally contains more condensed chromatin structures at the level of nucleosomes (Ricci et al., 2015). Consequently, primed ESCs might be less dependent on factors that additionally limit chromatin dynamics and flexibility, explaining why ESC+serum are less dependent on NoRC than ground state ESC+2i. Considering that expression of TIP5 and its genomic occupancy were similar between ESC+2i and ESC+serum, the results imply that NoRC function might be determined by the state of its substrate (i.e. chromatin), rather than NoRC itself. How NoRC mechanistically controls local chromatin dynamics remains elusive and will be an aim of our future investigations.

The ground state-specific role of TIP5 could be recapitulated by depleting or chemically inactivating TOP2A in ESC+2i and ESC+serum. Besides controlling the expression of TIP5-regulated genes, TOP2A was also required for correct H3K27me3 occupancy in ESC+2i without any detectable influence in ESC+serum. While cohesin and its contribution to 3D genome organization into TADs has been extensively studied in recent years (Rao et al., 2017; Schwarzer et al., 2017), the role of topoisomerases in regulating chromatin structures and genome architecture in mammalian cells remained under-investigated. Previous analyses of TOP2A in ESC+serum revealed an association with active regions that correlate poorly

with H3K27me3 (Dykhuizen et al., 2013; Thakurela et al., 2013), a binding profile that is similar to TIP5 genome occupancy found in our analyses. Unfortunately, due to the discontinuation of validated high-grade ChIP TOP2a antibodies (Naughton et al., 2013; Thakurela et al., 2013) we have been unable to measure TOP2A genome occupancy in ESC+2i. Computational modeling studies proposed an active role of topoisomerases and supercoiling in shaping the 3D genome structures (Benedetti et al., 2014a; Benedetti et al., 2014b). Accordingly, recent data indicated that transcription-induced supercoiling might act as the driving force of chromatin loop extrusion during the formation of TADs in interphase chromosomes (Racko et al., 2017). Moreover, TOP2B was found to associate with CTCF at TAD borders and was suggested to resolve torsional stress leading to the stable association of cohesin at TAD boundaries (Uuskula-Reimand et al., 2016). Of note is that inhibition of TOP2 activities or depletion of TOP2A or SMC3 resulted in similar changes in gene expression (upregulation or downregulation) as observed upon TIP5 KD in ESC+2i. These results point to a role of NoRC that, particularly in cooperation with TOP2A, might have a major function in shaping 3D genome architecture instead of directly acting as activator or repressor on gene expression.

The role of NoRC in genome architecture is further strengthened by the fact that TIP5 binds to large and active chromatin regions (type A compartment) and that its occupancy does not correlate with TIP5-regulated genes. We would like to further emphasize that the changes in gene expression and H3K27me3 do not seem to be directly coupled. Although differentially expressed genes were enriched in H3K27me3, TIP5 depletion in *Eed* and *Ring1b* KO in ESC+2i highly recapitulated gene expression changes observed in TIP5-depleted wt ESC+2i. Instead it seems that changes in gene expression and H3K27me3 content are two independent read outs of altered chromatin structures triggered by TIP5 depletion. Remarkably, genes affected by TIP5 knockdown are also often differentially expressed in ESC+2i and ESC+serum, suggesting that transcription of these genes – so far for unknown reasons – are in general more sensitive to changes in chromatin structures in ESCs.

The switch of NoRC function between ground state and primed ESCs suggests that NoRC action is highly dependent on the developmental and chromatin state of the cells. A change of NoRC function was also previously reported during the transition from pluripotency toward

differentiation and between normal and cancer cells. In somatic and healthy cells, TIP5 is exclusively localized within the nucleolus, the compartment where rRNA genes reside, and represses their transcription (Guetg et al., 2012; Santoro et al., 2002). By contrast, in both ground state and primed ESCs, the binding of TIP5 to rRNA genes and its nucleolar localization are impaired (this work and (Leone et al., 2017; Savic et al., 2014)). Similarly to ESCs, in metastatic prostate cancer (PCa) cells TIP5 is abundantly present in the nucleoplasm and regulates genes involved in developmental processes and frequently repressed in metastatic PCa (Gu et al., 2015b). Considering that cancer cells often acquire stem cell-like features, it is likely that NoRC uses similar mechanisms to regulate ESCs and cancer cells, a working hypothesis that is currently under investigation in our lab.

Our study showed the different responses triggered by loss of chromatin regulators between such closely related cell types (ground state and primed ESCs). These results suggest that the global shaping of chromatin is a fundamental player for gene regulation and chromatin states and highlight different degrees of genome plasticity during the early phases of development.

Materials and Method

Cell culture

One hundred and twenty-nine mouse embryonic stem cells (E14 line) were cultured in either 2i media composed of DMEM-F12 and Neurobasal medium (1:1, Life Technologies), supplemented with 1× N2/B27 (Life Technologies), 1× penicillin/streptomycin/l-glutamine (Life Technologies), 50 µM β-mercaptoethanol (Life Technologies), recombinant leukemia inhibitory factor, LIF (Polygene, 1,000 U/ml) and MEK and GSK3β inhibitors, 2i (Sigma CHIR99021 and PD0325901, 3 and 1 µM, respectively) or in serum medium containing DMEM (Life Technologies), 15% FCS, 1× MEM NEAA, 100 µM β-mercaptoethanol, 1× penicillin/streptomycin (Life Technologies). ESCs were seeded at a density of 50,000 cells/cm^2 in culture dishes (Corning® CellBIND® surface) coated with 0.1% gelatin without feeder layer. Propagation of cells was carried out every 2 days using enzymatic cell dissociation.

ESCs were differentiated by culturing for 48–72h in complete medium: DMEM, 10% FCS, 1 mM sodium pyruvate (Sigma), 1× NEAA (Life Technologies), 1× penicillin/streptomycin/l-glutamine, 100 µM β-mercaptoethanol on 0.1% gelatin-coated culture dishes.

Differentiation toward neural progenitor cells (NPCs) was obtained according to previously established protocol (Bibel et al., 2004). In brief, differentiation employed a suspension-based embryoid bodies formation (Bacteriological Petri Dishes, Bio-one with vents, Greiner®) in neural differentiation media (DMEM, 10% FCS, 1× MEM NEAA, 1× penicillin/streptomycin/l-glutamine, 100 µM β-mercaptoethanol). During the 8-day differentiation procedure, media were exchanged every 2 days. In the last 4 days of differentiation, the media were supplemented with 2µM retinoic acid (RA) to generate neuronal precursors. DNMT-TKO ESCs were kindly provided by M. Okano (Tsumura et al., 2006). *Eed*-KO and *Ring1b*-KO ESC+2i were a gift from C. Ciaudo (ETH, Zurich).

Transfections

ESCs were transfected with the indicated siRNAs (50 nM siRNA) using Lipofectamine® RNAiMAX (Life Technologies) in Opti-MEM® GlutaMAX™ (Life Technologies) reduced-serum medium. ESCs were seeded 2h prior to siRNA transfection and collected 3 days post

transfection if not stated differently. For analysis of differentiation potential of TIP5-depleted ESCs, transfected cells were re-seeded at equal cell numbers into complete differentiation medium 48h post transfection. Survival of differentiated cells was assessed 72h later. Efficiencies of siRNA-mediated depletions were monitored by qRT–PCR 3 days post-transfection.

Alkaline phosphatase (AP) staining

Cells were fixed in 4% paraformaldehyde for 10 min, washed with AP buffer (100 mM Tris-Cl pH 9.5, 100 mM NaCl, 50 mM MgCl2), and then incubated for 5–10 min in BCIP®/NBT liquid substrate system (Sigma). The reaction was blocked with 10 mM Tris and 1 mM EDTA for 10 min.

ICRF-193 treatment

ESCs were seeded 24h prior to treatment. ICRF-193 was added directly to the medium to final concentration of 500nM, as it has previously been reported for ESCs (Thakurela et al., 2013). Cells were harvested for respective experiments after 24h of inhibitor treatment.

RNA extraction, reverse transcription, and quantitative PCR (RT–qPCR)

RNA was purified with TRIzol reagent (Life Technologies). 1µg total RNA was primed with random hexamers and reverse-transcribed into cDNA using MultiScribe™ Reverse Transcriptase (Life Technologies). Amplification of samples without reverse transcriptase assured absence of genomic or plasmid DNA (data not shown). The relative transcription levels were determined by normalization to *Rps12* or *beta-Actin* mRNA levels, as indicated. qRT–PCR was performed with KAPA SYBR® FAST (Sigma) on a Rotor-Gene Q (Qiagen). Primer sequences are listed in **Supplementary Table 4**.

Chromatin fractionation

ESCs were collected by trypsinization, washed once with PBS and counted. ES cell pellets were resuspended at a concentration of 10mio cells/ml in chromatin fractionation buffer (10mM Hepes pH 7.6, 150mM NaCl, 3mM $MgCl_2$, 0.5% Triton X-100, 1mM DTT freshly

supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche)) and incubated for 30 minutes at room temperature rotating. Precipitated chromatin was fractionated by centrifugation. Total and chromatin fractionated samples were further processed by MNase (S7 Micrococcal nuclease, Roche) digest for ensuring sufficient genomic DNA fragmentation. All samples were incubated in 1x Laemmli buffer (10% glycerol, 10mM Tris pH 6.8, 2% SDS, 0.1mg/ml bromphenolblue, 2% β-mercaptoethanol) at 95°C for 5 minutes and were further analyzed by Western Blotting.

Generation of FLAG/HA-TIP5 ES cell line by CRISPR/Cas9

CRISPR/Cas9 cloning and targeting strategy was performed as previously described (Ran et al., 2013). sgRNA guide sequence (GTCGTTTGCCTCCATTTCTGT) was chosen to target the TIP5 locus on exon 3 three base pairs upstream of the ATG start codon and was cloned into pSpCas9(BB)-2A-GFP (PX458, Addgene). This plasmid was co-transfected with the HDR repair template plasmid containing the FLAG/HA inclusion flanked by 1kb homology arms into wild type ESCs at a molar ratio of 1:3. After two days, positively transfected cells were selected for GFP expression and were then further cultured for additional three days. Subsequently, ESCs were seeded for single cell clone isolation. Derived clones were genotyped by PCR using two different primer pairs. One PCR aimed to identify site-specific FLAG-HA inclusion, while the other PCR allowed distinguishing between inclusions in one or both alleles (as illustrated in Supplementary Fig.S4). The integrity of the FLAG-HA inclusion was verified by cloning the exon 3 sequences into CloneJET PCR cloning kit (Thermo Scientific) and by Sanger sequencing (Microsynth).

Chromatin immunoprecipitation and mass spectrometric analyses

Approximately 10^8 ESCs were collected by scraping followed by washing with PBS. Nuclei were isolated by re-suspending the cells in two consecutive rounds in hypotonic buffer (10mM Hepes pH 7.6, 1.5mM $MgCl_2$, 10mM KCl, 2mM Na_3VO_4 freshly supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche)). The suspension was passed through a douncer homogenizer with a loose pestle 10-20 times and the purity of nuclei was checked under a microscope. The chromatin was then isolated and crosslinked by resuspending the nuclei in

the chromatin fractionation/crosslinking buffer (10mM Hepes pH 7.6, 3mM $MgCl_2$, 150mM NaCl, 0.5% Triton X-100, 0.5mM dithiobis[succinimidylpropionate] (DSP, Thermo Scientific), 2mM Na_3VO_4 supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche)) and rotation at room temperature for 30min. The crosslinking was stopped by the addition of 25mM Tris HCl pH 7.5. The chromatin was then isolated by centrifugation and washed twice in MNase digestion buffer (0.3M Sucrose, 50mM Tris pH 7.5, 30mM KCl, 7.5mM NaCl, 4mM $MgCl_2$, 1mM $CaCl_2$, 0.125% NP-40, 0.25% NaDeoxycholate, 2mM Na_3VO_4 supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche)). Digestion of chromatin into mononucleosomes was assured by digestion with 100U MNase (Roche) in MNase digestion buffer at 37°C for 1h. SDS was then added to a final concentration of 1% followed by a 3x 30sec sonication steps with a bioruptor sonicator (Diagenode). Insoluble precipitates were removed by centrifugation and soluble crosslinked chromatin extracts were diluted 10x in IP buffer (0.3M Sucrose, 50mM Tris pH 7.5, 30mM KCl, 300mM NaCl, 4mM $MgCl_2$, 1mM $CaCl_2$, 0.125% NP-40, 0.25% NaDeoxycholate, 2mM Na_3VO_4 supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche)) and 30µl ANTI-FLAG M2 Affinity Gel (Sigma) were added to the extracts. Binding of FLAG/HA-TIP5 was performed by incubation over night at 4°C while rotating. The beads were subsequently washed five times in wash buffer (20mM Tris pH 7.5, 20% glycerol, 100mM KCl, 300mM NaCl, 1.5mM MgCl2, 0.2mM EDTA, 0.125% NP40, 0.25% NaDeoxycholate, 2mM Na_3VO_4 supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche)). Purified complexes were then eluted with 2mM FLAG peptide (Sigma) in TBS buffer (50mM Tris-HCl pH 8.0, 150mM NaCl). Eluted proteins were precipitated with the addition of 0.25x volume of 100% trichloric acid (Sigma). Protein pellets were washed five times with cold Acetone (Merck) and submitted for subsequent mass spectrometric analyses by the Functional Genomics Center Zurich (FGCZ). The dry pellets were dissolved in 45 µl buffer (10 mM Tris + 2 mM $CaCl_2$, pH 8.2) and 5 µl of trypsin (100 ng/µl in 10 mM HCl) for digestion, which was carried out in a microwave instrument (Discover System, CEM) for 30 min at 5 W and 60 °C. Samples were dried in a SpeedVac (Savant). For LC-MS/MS analysis the samples were dissolved in 0.1% formic acid (Romil)) and an aliquot ranging from 5 to 25% was analyzed on a nanoAcquity UPLC (Waters Inc.) connected to a Q Exactive mass spectrometer (Thermo Scientific) equipped with a Digital PicoView source

(New Objective). Peptides were trapped on a Symmetry C18 trap column (5 µm, 180 µm x 20 mm, Waters Inc.) and separated on a BEH300 C18 column (1.7 µm, 75 µm x 150 m, Waters Inc.) at a flow rate of 250 nl/min using a gradient from 1% solvent B (0.1% formic acid in acetonitrile, Romil)/99% solvent A (0.1% formic acid in water, Romil) to 40% solvent B/60% solvent A within 90 min. Mass spectrometer settings were: Data dependent analysis. Precursor scan range 350 – 1500 m/z, resolution 70'000, maximum injection time 100 ms, threshold 3e6. Fragment ion scan range 200 – 2000 m/z, Resolution 35'000, maximum injection time 120 ms, threshold 1e5. Proteins were identified using the Mascot search engine (Matrix Science, version 2.4.1). Mascot was set up to search the SwissProt database assuming the digestion enzyme trypsin. Mascot was searched with a fragment ion mass tolerance of 0.030 Da and a parent ion tolerance of 10.0 PPM. Oxidation of methionine was specified in Mascot as a variable modification. Scaffold (Proteome Software Inc.) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they achieved a false discovery rate (FDR) of less than 0.1% by the Scaffold Local FDR algorithm. Protein identifications were accepted if they achieved an FDR of less than 1.0% and contained at least 2 identified peptides.

Chromatin immunoprecipitation (ChIP)

ChIP analysis was performed as previously described (Leone et al., 2017). Briefly, 1% formaldehyde was added to cultured cells to cross-link proteins to DNA. For histone ChIPs, isolated nuclei were then lysed and sonicated using a Bioruptor ultrasonic cell disruptor (Diagenode) to shear genomic DNA to an average fragment size of 200 bp. 20 µg of chromatin was diluted to a total volume of 500 µl with ChIP buffer (16.7 mM Tris–HCl, pH 8.1, 167 mM NaCl, 1.2 mM EDTA, 0.01% SDS, 1.1% Triton X-100) and pre-cleared with 10 µl packed Sepharose beads for 2 h at 4°C. Pre-cleared chromatin was incubated overnight with the indicated antibodies. The next day, Dynabeads protein-A (or -G, Millipore) were added and incubated for 4 h at 4°C. After washing, bound chromatin was eluted with the elution buffer (1% SDS, 100 mM $NaHCO_3$). Upon proteinase K digestion (50°C for 3 h) and reversion of cross-linking (65°C, overnight), DNA was purified with phenol/chloroform, ethanol precipitated and quantified by qPCR using the primers listed in **Supplementary Table 4**.

For TIP5 ChIPs, we noticed that sonication of formaldehyde-crosslinked chromatin induced degradation of TIP5 (data not shown). Therefore, to increase the efficiency of TIP5 ChIPseq, we fragmented crosslinked chromatin into mono-nucleosomes through digestion with MNase. Briefly, isolated and crosslinked nuclei were MNase digested in 400µl MNase digestion buffer (0.3M Sucrose, 50mM Tris pH 7.5, 30mM KCl, 7.5mM NaCl, 4mM $MgCl_2$, 1mM $CaCl_2$, 0.125% NP-40, 0.25% NaDeoxycholate, 2mM Na_3VO_4 supplemented with cOmplete™ Protease Inhibitor Cocktail (Roche)) with 100U MNase (Roche) at 37°C for 1h. The digestion was then stopped with 5mM EDTA and the digested chromatin was solubilized in 1% SDS and three pulses of 30sec sonication using a Bioruptor ultrasonic cell disruptor (Diagenode). 200µg of pre-cleared chromatin was immunopurified with incubation of 30µl of Anti-FLAG M2 Affinity Gel (Sigma) or 30µl Anti-HA magnetic beads (Pierce) over night. The samples were subsequently washed, eluted and the DNA was purified as for histone ChIPs. ChIP-qPCR measurements were performed with KAPA SYBR® FAST (Sigma) on a Rotor-Gene Q (Qiagen) always comparing enrichments over input samples. Primer sequences are listed in **Supplementary Table 4**.

For ChIPseq analyses, the quantity and quality of the isolated DNA was determined with a **Qubit® (1.0) Fluorometer (Life Technologies,** California, USA) and a Bioanalyzer 2100 (Agilent, Waldbronn, Germany). The **Nugen Ovation Ultra Low Library Systems** (Nugen, Inc, California, USA) was used in the following steps. Briefly, ChIP samples (1 ng) was end-repaired and polyadenylated before the ligation of Illumina compatible adapters. The adapters contain the index for multiplexing. The quality and quantity of the enriched libraries were validated using **Qubit® (1.0) Fluorometer and the** Bioanalyzer 2100 (Agilent, Waldbronn, Germany)**. The libraries were normalized to 10nM in** Tris-Cl 10 mM, pH8.5 with 0.1% Tween 20. **The TruSeq SR Cluster Kit v4-cBot-HS** (Illumina, Inc, California, USA) was used for cluster generation using 8 pM of pooled normalized libraries on the cBOT. Sequencing was performed on the Illumina HiSeq 2500 single end 126 bp using the **TruSeq SBS Kit v4-HS** (Illumina, Inc, California, USA)**.**

ChIPseq data analysis

Own and published ChIPseq reads were aligned to the mouse mm10 reference genome using Bowtie2 (version 2.2.5; (Langmead and Salzberg, 2012)). Read counts were computed and normalized using “bamCoverage” from deepTools (version 2.0.1; (Ramirez et al., 2014)) using a bin size of 50bp. deepTools was also further used to generate all heat maps, profiles and pearson correlation plots. TIP5 bound regions were defined using SICER (version 1.1; (Zang et al., 2009)) by comparing the FLAG ChIPs of tagged vs non-tagged TIP5 ESCs in 2i and serum using the following arguments: W=1000 G=3000 FDR=0.00001. These analyses revealed 8824 and 7221 TIP5-bound regions in 2i and serum ESCs, respectively. Highly similar results were obtained by defining TIP5 bound regions comparing the FLAG ChIPs to the input samples, excluding strong biases of the FLAG antibody.

CTCF ChIPseq data sets in 2i and serum ESCs were taken from (Jonathan A. Beagan, 2017). CTCF peaks were defined using MACS2 (version 2.1.0; (Zhang et al., 2008)) comparing the CTCF ChIPseq to its respective input sample with a qValue cutoff of 0.0001. Using these parameters 56218 and 47245 CTCF peaks were defined in 2i and serum ESCs, respectively. The ratio of mean H3K27me3 read counts 2.5kb upstream and downstream of each CTCF peak were calculated and served as “insulation scores”. H3K27me3-insulated CTCF peaks were defined with an insulation score of >2 or <0.5. Regions with an overall average of less than 0.3 normalized read counts were excluded. This filtering resulted in 7547 CTCF peaks (13.4% of all CTCF peaks) and 8448 peaks (17.9% of all CTCF peaks) that were H3K27me3-insulated in 2i and serum ESCs, respectively. The accuracy of these calculations was confirmed by plotting published H3K27me3 ChIPseq data sets from wild type 2i and serum ESCs (Marks et al., 2012) over these H3K27me3-insulating CTCF peaks revealing highly similar results (data not shown). Pearson correlation plots were generated using deepTools (version 2.0.1). After removal of blacklist regions, chromosome 19 was partitioned into 1kb windows and correlation plots were computed with indicated data sets. H3K36me3 (GSM590119), H3K9me3 (GSM850407) (Marks et al., 2012) and H3K4me1 (GSM1856424) (Joshi et al., 2015) were taken from published ChIPseq data sets of ESCs in 2i. For data analysis over transcribed regions, genomic coordinates from all refseq transcripts were retrieved from Ensembl biomart. After removal of blacklist regions, normalized ChIPseq

data sets were plotted either over +/- 5kb from the TSS or over the entire transcribed region by scaling the gene length to 20kb (+5kb from TES and -5kb from the TSS). For distinguishing H3K27me3-high and H3K27me3-low promoters, mean read counts +/- 5kb from each TSS were computed. The first quartile was termed as H3K27me3-high, while quartiles 2-4 were defined as H3K27me3-low. TIP5-bound genes were defined by an overlap of a TIP5-bound region with the transcribed regions of the respective gene using bedtools (version 2.24.0; (Quinlan and Hall, 2010)). Conversion of mm9 and mm10 data sets was performed using Crossmap (version 0.2.4; (Zhao et al., 2014)). Integrative Genome Viewer (IGV, version 2.3.92) (Robinson et al., 2011) was used to visualize and extract representative ChIPseq tracks.

RNAseq and data analysis

Total RNA from three independent siRNA-mediated TIP5 knockdown experiments was purified with TRIzol reagent (Life Technologies) as stated above. In order to remove DNA contaminants, the samples were treated with 1U DNaseI (Thermo Scientific) for 1h at 37°C and the RNA samples were re-purified using TRIzol. The quality of the isolated RNA was determined with a **Qubit® (1.0) Fluorometer (Life Technologies,** California, USA) and a Bioanalyzer 2100 (Agilent, Waldbronn, Germany). Only those samples with a 260 nm/280 nm ratio between 1.8–2.1 and a 28S/18S ratio within 1.5–2 were further processed. The TruSeq RNA Sample Prep Kit v2 (Illumina, Inc, California, USA) was used in the succeeding steps. Briefly, total RNA samples (100-1000 ng) were poly A enriched and then reverse-transcribed into double-stranded cDNA. The cDNA samples was fragmented, end-repaired and polyadenylated before ligation of TruSeq adapters containing the index for multiplexing Fragments containing TruSeq adapters on both ends were selectively enriched with PCR. The quality and quantity of the enriched libraries were validated using **Qubit® (1.0) Fluorometer and the Caliper GX** LabChip® GX (**Caliper Life Sciences, Inc., USA).** The product is a smear with an average fragment size of approximately 260 bp. **The libraries were normalized to 10nM in** Tris-Cl 10 mM, pH8.5 with 0.1% Tween 20. **The TruSeq SR Cluster Kit HS4000** (Illumina, Inc, California, USA) was used for cluster generation using 10 pM of pooled normalized libraries on the cBOT. Sequencing was performed on the Illumina

HiSeq 4000 single end 100 bp using the **TruSeq SBS Kit HS4000** (Illumina, Inc, California, USA). Reads were aligned to the reference genome (ensembl version 82) with Subread (i.e. subjunc, version 1.4.6-p4; (Liao et al., 2013)) allowing up to 16 alignments per read (options: –trim5 10 –trim3 15 -n 20 -m 5 -B 16 -H –allJunctions). Count tables were generated with Rcount (Schmid and Grossniklaus, 2015) with an allocation distance of 100 bp for calculating the weights of the reads with multiple alignments, considering the strand information, and a minimal number of 5 hits. Variation in gene expression was analyzed with a general linear model in R with the package edgeR (version 3.12.0; (Robinson and Oshlack, 2010)) according to a crossed factorial design with two explanatory factors (i) siRNA against Tip5 and a mock sequence and (ii) ESCs grown in 2i or serum. Genes differentially expressed between specific conditions were identified with linear contrasts using **trended** dispersion estimates and Benjamini-Hochberg multiple testing corrections. Genes with a *P*-value below **0.05** and a minimal fold change of 1.5 were considered to be differentially expressed. These thresholds have previously been used characterizing chromatin remodeler functions (de Dieuleveult et al., 2016). Gene ontology analysis was performed with David Bionformatics Resource 6.8 (Huang et al., 2009).

HiC and data analysis

HiC experiments were performed in triplicates of ESC+2i treated with siRNA-control or siRNA-*TIP5*. Five million cells were pelleted and resuspended in PBS-10%FCS. PBS-10%FCS-4% formaldehyde was added to a final concentration of 2% formaldehyde (v/v). Samples were incubated at room temperature for 10 minutes with mixing. Ice-cold glycine solution was added to a final concentration of 0.2M and immediately centrifuged for 5 minutes at 300xg at 4°C. Cells were washed in 1 ml of ice-cold PBS. Pellet was flash-frozen in liquid nitrogen. Pellet was taken up and washed in 1 ml ice cold lysis buffer1, resupended again in 1 ml of ice cold lysis buffer and incubated for 30 minutes at 4°C. Sample was pelleted and washed in 0.5 ml 1.2x DpnII buffer, resuspended in 0.5 ml of DpnII buffer again and moved to a thermomixer at 37°C and 300 rpm. SDS was then carefully added to a concentration of 0.3%, slowly suspended with a pipet and incubated for an hour. Triton X-100 was added to a concentration of 2.6% and sample was incubated for another hour. To digest the sample, 200

units of DpnII enzyme were added for a 4 hour incubation in a thermomixer at 37°C and 900 rpm; another 200 units of DpnII were added for overnight incubation. From here on, we adopted the protocol described in (Rao et al., 2014) with some adjustments Cells were incubated for 20 minutes at 65°C to heat inactivate DpnII, pelleted and 300 µl of fill-in mastermix was added (218 µl of MilliQ, 30 µl of 10x NEB buffer 2, 15 µl of 10mM dCTP, 15 µl of 10mM dGTP, 15 µl of 10mM dTTP, 37.5 µl of 0.4mM biotin-14-dATP (Life Technologies, 19524-016), 10 µl of 5U/µl DNA Polymerase I, Large (Klenow) Fragment (NEB, M0210)). Sample was mixed by pipetting and incubated for 60 minutes at 37°C shaking at 300 rpm and placed at 4°C afterwards. 900 µl of ligation mix was added (120 µl 10x ligase buffer, 50 units of T4 ligase (Roche) and 770 µl MilliQ) and mixed by inverting and incubated overnight at 16°C. Sample was pelleted for 5 minutes at 1000 x g and taken up in 500 µl 10 mM Tris. Protein was degraded by adding 50 µl of 20mg/ml proteinase K (NEB, P8102), 50 µl of 10% SDS and incubated at 55°C for 30 minutes. 57 µl of 5M of sodium chloride was then added and the sample was incubated at 68°C overnight or for at least 1.5 hours. Samples were cooled to room temperature and DNA was purified using NucleoMag P-Beads and taken up in 5mM Tris pH7.5. Samples were sheared to a size of 300-500 bp using a Covaris S2 focused-ultrasonicator. From here on, the protocol described in (Rao et al., 2014) was adopted.
FastQ files were mapped to the mouse genome (mm10) using bwa-mem (Li and Durbin, 2010) and filtered and deduplicated using HiCUP v0.5.10 (Wingett et al., 2015). Chromosomal interaction matrices were generated using Juicer (Durand et al., 2016) at 500 Kb resolution and normalized by Knight and Ruiz's matrix balancing algorithm. Biological replicates were first processed independently and inspected for clustering between TIP5 depletion and control conditions by PCA. Next, replicates per condition were pooled to create merged contact maps that were used in the downstream analyses.

To visualize the impact of Tip5 knockdown in chromosomal architecture, we plotted the median contact frequency from each genomic region at increasing genomic distance. ENCODE Data Analysis Consortium Blacklisted Regions (Hoffman et al., 2013) were excluded from the analysis.

Public Datasets Used in This Study

Public datasets used in this study are listed in Table 4.

Accession numbers

All raw data generated in this study using high throughput sequencing are accessible through NCBI's GEO (accession number GSE112222).

Acknowledgements

This work was supported by the Swiss National Science Foundation (310003A-152854 and 31003A_173056 to R.S.; PP00P3_150667 and NCCR RNA & Disease to ACM), Forschungskredit of the University of Zurich (to D.D and E.V.), Julius Müller Stiftung, Olga Mayenfisch Stifung and Stiftung für wissenschaftliche Forschung an der Universität Zürich (to R.S.). We thank Peter Hunziker, Catherine Aquino and the Functional Genomic Center Zurich for the assistance in sequencing and proteomic analysis. We also thank Dominik Bär for technical assistance and Rostyslav Kuzyakiv from S3IT of the University of Zurich for assistance in bioinformatic analyses. We thank C. Ciaudo for having provided ESC lines.

Figure legends

Figure 1

TIP5 is required for proliferation and differentiation of ESC+2i

A. TIP5 expression is higher in ESC+2i than in differentiated cells (neural progenitors, NPC). Left panel. *TIP5* mRNA levels were measured by qRT-PCR and normalized to *Rps12* mRNA and to ESC+2i. Average values of three independent experiments. Error bars represent s.d. Right panel. Western blot showing TIP5 protein levels in ESC+2i and NPC. Tubulin is shown as a protein loading control.

B. TIP5 associates with chromatin of ESC and NPC. Chromatin-bound (Chrom.) and soluble (Sol.) fractions of equivalent cell number of ESC+2i and NPCs were analyzed by western blot for TIP5 levels. Tot., total protein. Tubulin and histones are shown as loading and fractionation control.

C. TIP5 is expressed at similar levels in both ESC+2i and ESC+serum. Left panel. *TIP5* mRNA levels were measured by qRT-PCR and normalized to *Rps12* mRNA and to ESC+2i. Average values of three independent experiments. Error bars represent s.d.. Right panel. Western blot showing TIP5 protein levels in ESC+2i and ESC+serum. SNF2H is shown as a protein loading control.

D. siRNA-knockdown efficiency of TIP5 shown by qRT-PCR and western blot. *TIP5* mRNA levels were measured by qRT-PCR and normalized to *Rps12* mRNA and to each ESC line. Average values of three independent experiments. Error bars represent s.d. Right panel. Western blot showing TIP5 protein levels. Tubulin is shown as a protein loading control.

E. TIP5 knockdown affects proliferation of ESC+2i but not of ESC+serum. Data represent relative cell numbers after 3 days of siRNA treatment and were normalized to ESC transfected with siRNA-Control. Average values of three independent experiments. Error bars represent s.d. Statistical significance (*P*-values) for the experiments was calculated using the paired two-tailed t-test (** < 0.01; ns, non-significant).

F. TIP5 is required for the differentiation of ESC+2i but not of ESC+serum. Representative images of alkaline phosphatase staining of ESC and cells after 3 days of differentiation.

G. Quantification of differentiated cells. Values represent relative number of differentiated cells from three independent experiments relative to control cells.

Figure 2

TIP5 regulates expression of bivalent genes in ESC+2i

A. Volcano plot showing fold change (log_2 values) in transcript level of ESC+2i and ESC+serum upon TIP5 knockdown. Gene expression values of three replicates were averaged and selected for 1.5 fold changes and P <0.05.

B. Top 10 biological process gene ontology terms as determined using DAVID for genes regulated by TIP5 in ESC+2i and ESC+serum.

C. Left panel. Heat map profiles of H3K4me3 and H3K27me3 at ±5 kb from the TSS of genes differentially regulated by TIP5 in ESC+2i and random genes. Differentially regulated genes are shown as upregulated or downregulated in ESC+2i upon TIP5 knockdown. Data were ranked by H3K27me3 levels. Right panel. Average density profiles of H3K4me3, H3K27me3 and H3K36me3 at genes upregulated or downregulated in ESC+2i upon TIP5 knockdown and random genes. Transcription start site (TSS), transcription end site (TES).

D. Example of transcription levels and H3K4me3 and H3K27me3 occupancy at a bivalent gene upregulated (*Efnb2*) and downregulated (*Rasip1*) upon TIP5 knockdown in ESC+2i but not in ESC+serum.

E. Genes regulated by TIP5 in ESC+2i are expressed at low level. Transcription levels are expressed as normalized RPKM. Box plot with median bar, first-third quartile box and 5th–95th percentile whiskers. (*** <0.001, two-tailed t-test.)

F. Validation by qRT-PCR of genes regulated by TIP5 in ESC+2i but not in ESC+serum. *Nanog*, *Rex1* and *Actin* B (*ActB*) are shown as genes not regulated by TIP5. mRNA levels were normalized to *Rps12* mRNA and to ESCs transfected with siRNA-Control. Average values of three independent experiments. Error bars represent s.d.

Figure 3

TIP5 regulates H3K27me3 occupancy in ESC+2i

A. Western blot showing total levels of H3K27me3, H3K27ac and H3K4me3 in ESC+2i transfected with siRNA-control and siRNA-*Tip5*. Histone H3 is shown as a protein loading control.

B. Average density plot of ChIPseq read counts of H3K4me3, H3K27ac and H3K27me3 at -/+ 5 kb from the TSS of refseq genes in ESCs treated with siRNA-control or siRNA-*Tip5*. Q1: first quartile corresponding to sequences with high H3K27me3 content. Q2-4: quartiles 2-4 corresponding to sequences with low H3K27me3 content.

C. ChIP-qPCR of H3K27me3 in ESC+2i and ESC+serum treated with siRNA-control or siRNA-*Tip5.* Data were measured by qPCR and normalized to input and to *Tshz1* value in control cells. Control represents an intergenic sequence that does not contain H3K27me3. Average values of three independent experiments. Error bars represent s.d.

D. Average density and heatmap profile of H3K27me3 transition state regions marked by CTCF sites in ESC+2i or ESC+serum transfected with siRNA-control or siRNA-*Tip5*. Data were ranked by H3K27me3 content in the corresponding siRNA-control ESCs. Cluster 1 represents high to low H327me3 transition whereas Cluster 2 shows low to high H3K27me3 levels.

E. H3K27me3 ChIP-qPCR in ESC+2i (top panel) or ESC+serum (bottom panel) transfected with siRNA-control or siRNA-*Tip5*. Values show H3K27me3 occupancy at four sequences (#1-#4) with elevated H3K27me3 at CTCF boundary (H3K27me3^{hi}/CTCF). Data were measured by qPCR and normalized to input and to sequence #1 value in control cells. Average values of three independent experiments. Error bars represent s.d.

Figure 4

TIP5 associates with large and active chromatin domains in ESC+2i and ESC+serum

A. Western blot showing equal TIP5 levels in wt-ESC and F/H-TIP5 ESC line containing FLAG-HA sequences at the N-terminus of both *Tip5* alleles. Whole cell lysates from equivalent amounts of cells were analysed. SNF2H and Tubulin are shown as protein loading controls.

B. Representative images showing the association of TIP5 at regions enriched in H3K27ac and low in H3K27me3. The grey rectangles highlighted some of TIP5 associated regions. The data also show the decrease of H3K27me3 levels in ESC+2i upon TIP5 knockdown at regions not bound by TIP5.

C. Boxplot showing the mean of normalized TIP5 occupancies in ESC+2i and ESC+serum at TIP5-bound regions. Statistical significance (*P*-values) for the experiments was calculated using the unpaired two-tailed t-test (**** < 0.0001).

D. Anti-FLAG ChIP of wt-ESCs and F/H-TIP5-ESCs cultured in 2i or serum. Data were measured by qPCR and normalized to input and a control region that is not bound by TIP5. Average values of three independent experiments. Error bars represent s.d.

E. Pearson correlation heat map for the indicated ChIPseqs. TIP5 ChIPseqs were performed with FLAG antibodies in *F/H*-TIP5-ESCs cultured in 2i or serum and with HA antibodies in *F/H*-TIP5-ESC+2i. Data of H3K4me3, H3K27ac and H3K27me3 are from this work. H3K36me3, H3K9me3 and H3K4me1 values in ESC+2i were taken from published data sets (Joshi et al., 2015; Marks et al., 2012).

F. Average density profiles of ChIPseq read counts of H3K4me3, H3K27ac and H3K27me3 in ESCs at -/+ 5 kb from the TSS of genes bound by TIP5 and random genes.

G. Average density profiles of H3K4me3, H3K27ac and H3K27me3 at TIP5 bound refseq genes in ESCs treated with siRNA-control or siRNA-*Tip5.*

Figure 5

TIP5 associates with TOP2A and component of the cohesin complex on chromatin.

A. Scheme representing the strategy used to identify protein interacting with TIP5 on chromatin (Chromatin-IP).

B. TIP5-interacting proteins on chromatin. Mass spectrometry analysis of FLAG immunoprecipitates from wt and F/H-TIP5 ESC+2i. Values of peptide number are shown as the difference between peptides obtained in FLAG-IP of F/H-TIP5 and wt ESC chromatin extracts. Data are from three independent experiments. The proteins found enriched in FLAG-IP of F/H-TIP5 ESCs in all three experiments are shown. Further TIP5-interacting proteins identified in only two or one of the immunoprecipitation (IP) experiments are listed in Supplementary Table 3.

C. Anti-FLAG IP of chromatin purified from wt- and F/H-TIP5-ESCs. Western blot shows the interaction of TIP5 with SNF2H, TOP2A and SMC1a. After protein transfer, membrane was

first incubated with anti-SMC1a antibodies and after signal detection re-incubated with anti-TOP2A antibodies.

Figure 6

TOP2A and cohesin interact with TIP5 to regulate gene expression and H3K27me3 occupancy of ESC+2i

A. TOP2A and TIP5 regulate gene expression in ESC+2i via a shared pathway. qRT-PCR of genes regulated by TIP5 in ESC+2i (left panel) and in ESC+serum (right panel) upon TOP2A or TIP5 knockdown by siRNA. *Nanog*, *Rex1* and *Actin* B (*ActB*) are shown as genes not regulated by TIP5. mRNA levels were normalized to *Rps12* mRNA and to ESCs transfected with siRNA-Control. Average values of three independent experiments. Error bars represent s.d.

B. Inhibition of type 2 topoisomerase activity phenocopies the alterations in gene expression of TIP5-regulated genes in ESC+2i observed upon TOP2A or TIP5 knockdown. ESCs were treated with the TOP2 inhibitor ICF-193. mRNA levels were normalized to *Rps12* mRNA and to ESCs treated with DMSO. Average values of three independent experiments. Error bars represent s.d.

C. SMC3 and TIP5 regulate gene expression in ESC+2i via a shared pathway. qRT-PCR of genes regulated by TIP5 in ESC+2i (left panel) and in ESC+serum (right panel) upon SMC3 or TIP5 knockdown by siRNA. mRNA levels were normalized to *Rps12* mRNA and to ESC transfected with siRNA-Control. Average values of three independent experiments. Error bars represent s.d.

D,E. ChIP analysis showing that TOP2A regulates H3K27me3 occupancy in ESC+2i (D) but not in ESC+serum (E). Data were measured by qPCR and normalized to input and to *Tshz1* value in cells transfected with siRNA-control. Control represents an intergenic region that does not contain H3K27me3. Average values of three independent experiments. Error bars represent s.d.

F. Depletion of Smc3 does not affect H3K27me3 occupancy in ESC+2i. H3K27me3 ChIP values represent the average values of three independent experiments. Error bars represent s.d.

G,H. TIP5 binding in ESC+2i depends on TOP2 activity. (**G**) Anti-FLAG ChIP in F/H-TIP5 ESCs treated with siRNA-control or siRNA-*Top2a*. (**H**) Anti-FLAG ChIP in F/H-TIP5-ESCs cultured in 2i treated for 24 hours with DMSO or ICF-193. Values were normalized to input and to ATF7IP in control cells and are the average values of three independent experiments. Error bars represent s.d.

I. TIP5 binding in ESC+2i depends on cohesin. Anti-FLAG ChIP in F/H-TIP5-ESCs cultured in 2i treated with siRNA-control or siRNA-*Smc3*. Values were normalized to input and to ATF7IP in control cells and are the average values of three independent experiments. Error bars represent s.d.

Figure 7. TIP5 regulates genome architecture

A. Number of TIP5-bound regions in A and B compartments, respectively.

B. Relative contact frequencies over genomic distances in ESCs treated with siRNA-control and siRNA-TIP5.

C.D. Representative images from *HoxA* (**C**) and *HoxB* (**D**) gene clusters. Blue arrows indicated gain or increase of contacts in ESC+siRNA-*Tip5* compared to control cells (siRNA-control). Orange arrows indicate loss or decrease in contacts. HiC contacts were visualized using Juicebox (Rao et al., 2014)

Supplementary Figure legends

Supplementary Figure 1

TIP5 knockdown affects proliferation of ESC+2i

A. Data represent relative cell numbers and were normalized to ESC transfected with siRNA-Control. TIP5 knockdown was achieved with a different siRNA-*Tip5* sequence (siRNA-*Tip5#2*). Average values of three independent experiments. Error bars represent s.d.

B. Data represent relative cell numbers using another ESC line (ESC#2) cultured in 2i medium and were normalized to ESC transfected with siRNA-Control. Average values of two independent experiments.

Supplementary Figure 2

TIP5 regulates expression of bivalent genes in ESC+2i

A. Venn diagrams showing number of differently expressed genes upon TIP5 knockdown detected in ESC+2i compared to ESC+serum.

B. Venn diagrams showing genes differentially expressed in ESC+2i and in ESC+serum compared to genes regulated by TIP5 in ESC+2i.

C. Average density profiles from published ChIPseq data sets at genes upregulated or downregulated upon TIP5 depletion and at random genes in ESC+2i. ChIP data for Suz12 and Ring1b were from (Joshi et al., 2015) whereas EZH2 and Pol II were from (Marks et al., 2012).

D. qRT-PCR. Expression analysis of TIP5 regulated genes in ESC+2i upon TIP5 knockdown using a different siRNA-*Tip5* (siRNA-*Tip5#2*). mRNA levels were normalized to *Rps12* mRNA and to ESC+2i transfected with siRNA-Control. Average values of three independent experiments. Error bars represent s.d.

E. qRT-PCR. Expression analysis of TIP5 regulated genes in another ESC line (ESC#2) cultured in 2i medium upon TIP5 knockdown. mRNA levels were normalized to *Rps12* mRNA and to ESC+2i transfected with siRNA-Control. Average values of two independent experiments.

Supplementary Figure 3

A. Western blot showing unchanged levels of total H3K27me3 in ESC+2i upon TIP5 knockdown. Whole cell lysates from equivalent amounts of cells were analysed. Histone H3 is shown as a protein loading control.

B. Heat map profiles for H3K4me3, H3K27ac and H3K27me3 ChIPseq experiments at all gene promoters in ESC+2i (H3K4me3 and H3K27ac) and in both ESC+2i and ESC+serum (H3K27me3) treated with siRNA-control or siRNA-*Tip5*. Heat maps for H3K27me3 were further subdivided into H3K27me3-high (Q1) and H3K27me3-low (Q2-4) promoters. Data were individually ranked by H3K4me3, H3K27ac and H3K27me3 in the corresponding control cells (siRNA-Control). Q1: first quartile corresponding to sequences with high H3K27me3 content. Q2-4: quartiles 2-4 corresponding to sequences with low H3K27me3 content.

C. H3K27me3 ChIP in ESC+2i upon TIP5 knockdown using a different siRNA-*Tip5* (siRNA-*Tip5#2*). Values show H3K27me3 occupancy at promoters of genes regulated by TIP5 and at four sequences (#1-#4) with elevated H3K27me3 at CTCF boundaries ($H3K27me3^{hi}$/CTCF). Data were measured by qPCR and normalized to input and *Tshz1* value in control cells. Average values of three independent experiments. Error bars represent s.d.

D. TIP5 knockdown does not affect proliferation of TKO-ESCs. Data represent relative cell numbers after 3 days of siRNA treatment and were normalized to ESC transfected with siRNA-Control. Average values of three independent experiments. Error bars represent s.d.

E. Knockdown of TIP5 in TKO-ESCs does not affect transcription of genes regulated by TIP5 in ESC+2i. mRNA levels were normalized to *Rps12* mRNA and to ESCs transfected with siRNA-Control. Average values of three independent experiments. Error bars represent s.d.

F. ChIP of H3K27me3 in TKO-ESCs treated with siRNA-control or siRNA-*Tip5.* Data were measured by qPCR and normalized to input and to *Tshz1* value in control cells. Control represents an intergenic region that does not contain H3K27me3. Average values of three independent experiments. Error bars represent s.d.

G. TIP5 knockdown affect proliferation of $Eed^{-/-}$ and $Ring1b^{-/-}$ ESCs cultured in 2i. Data represent relative cell numbers after 3 days of siRNA treatment and were normalized to ESC transfected with siRNA-Control. Average values of three independent experiments. Error bars represent s.d.

H, I. TIP5 depletion also affects gene expression in the absence of PRC-mediated repression. RT-qPCR experiments in $Eed^{-/-}$ (**H**) and $Ring1b^{-/-}$ (**I**) -ESCs cultured in 2i treated with siRNA-control or siRNA-*Tip5.* mRNA levels were normalized to *Rps12* mRNA and ESC+2i transfected with siRNA-Control. Average values of three independent experiments. Error bars represent s.d.

Supplementary Figure 4

A. Scheme representing the strategy to insert FLAG-HA sequences at the 5' of the endogenous TIP5 sequence.

B, C. PCR genotyping of four clones for insertion of the F/H-TIP5 sequence on the *TIP5* locus. Clones #1 and #3 were homozygous for *F/H-TIP5* alleles. Sanger sequencing confirmed the integrity of F/H-TIP5.

D. Venn diagrams showing TIP5 bound regions defined by SICER in ESC+2i compared to ESC+serum.

E. Representative image showing that TIP5 associates with internal chromatin domains (ICD) in ESC+2i and ESC+serum. The data refer to TIP5 ChIPseq performed with FLAG-antibodies with *F/H-Tip5* ESC+2i and ESC+serum. Data for LADs and ICDs refer to ESC+serum and are from (Peric-Hupkes et al., 2010).

F. Venn diagrams showing genes bound by TIP5 compared to TIP5-regulated or random genes in ESC+2i.

Supplementary Figure 5

A. Correlation scores of PCA analysis defining A and B compartments from HiC experiments in siRNA-control and siRNA-*Tip5* treated ESC+2i.

B. Chromosome 1 as representative example for the global increase in far-*cis* contacts upon TIP5 depletion in ESC+2i. Differences in HiC contact frequencies (siRNA-*Tip5* - siRNA-Control) were visualized using Juicebox (Rao et al., 2014)

C, D. Representative images indicating changes in contact frequencies. Blue arrows show gain or increase of contacts in ESC+siRNA-*Tip5* compared to control cells (siRNA-control).

Orange arrows indicate loss or decrease in contacts. HiC contacts were visualized using juicebox (Rao et al., 2014)

Reference

Akimitsu, N., Adachi, N., Hirai, H., Hossain, M.S., Hamamoto, H., Kobayashi, M., Aratani, Y., Koyama, H., and Sekimizu, K. (2003). Enforced cytokinesis without complete nuclear division in embryonic cells depleting the activity of DNA topoisomerase IIalpha. Genes Cells *8*, 393-402.

Alexandrova, S., Kalkan, T., Humphreys, P., Riddell, A., Scognamiglio, R., Trumpp, A., and Nichols, J. (2016). Selection and dynamics of embryonic stem cell integration into early mouse embryos. Development *143*, 24-34.

Beagan, J.A., Duong, M.T., Titus, K.R., Zhou, L., Cao, Z., Ma, J., Lachanski, C.V., Gillis, D.R., and Phillips-Cremins, J.E. (2017). YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. Genome Res *27*, 1139-1152.

Benedetti, F., Dorier, J., Burnier, Y., and Stasiak, A. (2014a). Models that include supercoiling of topological domains reproduce several known features of interphase chromosomes. Nucleic acids research *42*, 2848-2855.

Benedetti, F., Dorier, J., and Stasiak, A. (2014b). Effects of supercoiling on enhancer-promoter contacts. Nucleic acids research *42*, 10425-10432.

Bibel, M., Richter, J., Schrenk, K., Tucker, K.L., Staiger, V., Korte, M., Goetz, M., and Barde, Y.A. (2004). Differentiation of mouse embryonic stem cells into a defined neuronal lineage. Nat Neurosci *7*, 1003-1009.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. Nature reviews Genetics *17*, 661-678.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nature cell biology *16*, 516-528.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu Rev Biochem *78*, 273-304.

de Dieuleveult, M., Yen, K., Hmitou, I., Depaux, A., Boussouar, F., Dargham, D.B., Jounier, S., Humbertclaude, H., Ribierre, F., Baulard, C.*, et al.* (2016). Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. Nature.

de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M.*, et al.* (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227-231.

de Wit, E., Vos, E.S., Holwerda, S.J., Valdes-Quezada, C., Verstegen, M.J., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H., and de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. Molecular cell *60*, 676-684.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell systems *3*, 95-98.

Dykhuizen, E.C., Hargreaves, D.C., Miller, E.L., Cui, K., Korshunov, A., Kool, M., Pfister, S., Cho, Y.J., Zhao, K., and Crabtree, G.R. (2013). BAF complexes facilitate decatenation of DNA by topoisomerase IIalpha. Nature *497*, 624-627.

Ficz, G., Hore, T.A., Santos, F., Lee, H.J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y.L., Walter, J.*, et al.* (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. Cell Stem Cell *13*, 351-359.

Galonska, C., Ziller, M.J., Karnik, R., and Meissner, A. (2015). Ground State Conditions Induce Rapid Reorganization of Core Pluripotency Factor Binding before Global Epigenetic Reprogramming. Cell Stem Cell *17*, 462-470.

Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. Nature reviews Molecular cell biology *12*, 36-47.

Gorkin, D.U., Leung, D., and Ren, B. (2014). The 3D genome in transcriptional regulation and pluripotency. Cell Stem Cell *14*, 762-775.

Gu, L., Frommel, S.C., Oakes, C.C., Simon, R., Grupp, K., Gerig, C.Y., Bar, D., Robinson, M.D., Baer, C., Weiss, M.*, et al.* (2015). BAZ2A (TIP5) is involved in epigenetic alterations in prostate cancer and its overexpression predicts disease recurrence. Nature genetics *47*, 22-30.

Guetg, C., Lienemann, P., Sirri, V., Grummt, I., Hernandez-Verdun, D., Hottiger, M.O., Fussenegger, M., and Santoro, R. (2010). The NoRC complex mediates the heterochromatin formation and stability of silent rRNA genes and centromeric repeats. The EMBO journal *29*, 2135-2146.

Guetg, C., Scheifele, F., Rosenthal, F., Hottiger, M.O., and Santoro, R. (2012). Inheritance of silent rDNA chromatin is mediated by PARP1 via noncoding RNA. Molecular cell *45*, 790-800.

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y.*, et al.* (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. Cell *162*, 900-910.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yanez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B.*, et al.* (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell *169*, 693-707 e614.

Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C.*, et al.* (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell stem cell *13*, 360-369.

Hackett, J.A., and Surani, M.A. (2014). Regulatory principles of pluripotency: from the ground state up. Cell stem cell *15*, 416-430.

Hoffman, M.M., Ernst, J., Wilder, S.P., Kundaje, A., Harris, R.S., Libbrecht, M., Giardine, B., Ellenbogen, P.M., Bilmes, J.A., Birney, E.*, et al.* (2013). Integrative annotation of chromatin elements from ENCODE data. Nucleic acids research *41*, 827-841.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Jonathan A. Beagan, M.T.D., Katelyn R. Titus, Linda Zhou, Zhendong Cao, Jingjing Ma, Caroline V. Lachanski, Daniel R. Gillis, Jennifer E. Phillips-Cremins (2017). YY1 and CTCF orchestrate a 3-D chromatin looping switch during early neural lineage commitment. Genome Research.

Joshi, O., Wang, S.Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P.J., Habibi, E., Shaik, J., Saeed, S., Handoko, L.*, et al.* (2015). Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. Cell stem cell *17*, 748-757.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A.*, et al.* (2013). Naive pluripotency is associated with global DNA hypomethylation. Nat Struct Mol Biol *20*, 311-316.

Leone, S., Bar, D., Slabber, C.F., Dalcher, D., and Santoro, R. (2017). The RNA helicase DHX9 establishes nucleolar heterochromatin, and this activity is required for embryonic stem cell differentiation. EMBO reports *18*, 1248-1262.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589-595.

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic acids research *41*, e108.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A.*, et al.* (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell *149*, 590-604.

Mayer, C., Schmitz, K.M., Li, J., Grummt, I., and Santoro, R. (2006). Intergenic transcripts regulate the epigenetic state of rRNA genes. Mol Cell *22*, 351-361.

Merkenschlager, M., and Nora, E.P. (2016). CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. Annu Rev Genomics Hum Genet *17*, 17-43.

Naughton, C., Avlonitis, N., Corless, S., Prendergast, J.G., Mati, I.K., Eijk, P.P., Cockroft, S.L., Bradley, M., Ylstra, B., and Gilbert, N. (2013). Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. Nat Struct Mol Biol *20*, 387-395.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell stem cell *4*, 487-492.

Nitiss, J.L. (2009). DNA topoisomerase II and its growing repertoire of biological functions. Nat Rev Cancer *9*, 327-337.

Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944 e922.

Nozaki, T., Imai, R., Tanbo, M., Nagashima, R., Tamura, S., Tani, T., Joti, Y., Tomita, M., Hibino, K., Kanemaki, M.T.*, et al.* (2017). Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. Molecular cell *67*, 282-293 e287.

Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W., Solovei, I., Brugman, W., Graf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M.*, et al.* (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. Molecular cell *38*, 603-613.

Pommier, Y., Leo, E., Zhang, H., and Marchand, C. (2010). DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. Chem Biol *17*, 421-433.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Racko, D., Benedetti, F., Dorier, J., and Stasiak, A. (2017). Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes. Nucleic acids research.

Ralf Strohner, R.S., Ingrid Grummt (2001). NoRC-a novel member of mammalian ISWI-containing chromatin remodeling machines. The EMBO journal.

Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research *42*, W187-191.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nature protocols *8*, 2281-2308.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D.*, et al.* (2017). Cohesin Loss Eliminates All Loop Domains. Cell *171*, 305-320 e324.

Ricci, M.A., Manzo, C., Garcia-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell *160*, 1145-1158.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature Biotechnology.

Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology.

Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J.*, et al.* (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proceedings of the National Academy of Sciences of the United States of America *112*, E6456-6465.

Santoro, R., Li, J., and Grummt, I. (2002). The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. Nature genetics *32*, 393-396.

Savic, N., Bar, D., Leone, S., Frommel, S.C., Weber, F.A., Vollenweider, E., Ferrari, E., Ziegler, U., Kaech, A., Shakhova, O.*, et al.* (2014). lncRNA Maturation to Initiate Heterochromatin Formation in the Nucleolus Is Required for Exit from Pluripotency in ESCs. Cell stem cell *15*, 720-734.

Savić, N., Bär, D., Leone, S., Frommel, S.C., Weber, F.A., Vollenweider, E., Ferrari, E., Ziegler, U., Kaech, A., Shakhova, O.*, et al.* (2014). lncRNA Maturation to Initiate Heterochromatin Formation in the Nucleolus Is Required for Exit from Pluripotency in ESCs. In Cell Stem Cell, pp. 720-734.

Schlesinger, S., Selig, S., Bergman, Y., and Cedar, H. (2009). Allelic inactivation of rDNA loci. Genes & development *23*, 2437-2447.

Schmid, M.W., and Grossniklaus, U. (2015). Rcount: simple and flexible RNA-Seq read counting. Bioinformatics *31*, 436-437.

Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M.*, et al.* (2015). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. Nat Genet *47*, 1179-1186.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., C, H.H., Mirny, L.*, et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51-56.

Thakurela, S., Garding, A., Jung, J., Schubeler, D., Burger, L., and Tiwari, V.K. (2013). Gene regulation and priming by topoisomerase IIalpha in embryonic stem cells. Nature communications *4*, 2478.

Tiwari, V.K., Burger, L., Nikoletopoulou, V., Deogracias, R., Thakurela, S., Wirbelauer, C., Kaut, J., Terranova, R., Hoerner, L., Mielke, C.*, et al.* (2012). Target genes of Topoisomerase IIbeta regulate neuronal survival and are defined by their chromatin state. Proc Natl Acad Sci U S A *109*, E934-943.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R.*, et al.* (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes to cells : devoted to molecular & cellular mechanisms *11*, 805-814.

Uuskula-Reimand, L., Hou, H., Samavarchi-Tehrani, P., Rudan, M.V., Liang, M., Medina-Rivera, A., Mohammed, H., Schmidt, D., Schwalie, P., Young, E.J.*, et al.* (2016). Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. Genome Biol *17*, 182.

Voigt, P., Tee, W.W., and Reinberg, D. (2013). A double take on bivalent promoters. Genes & development *27*, 1318-1338.

Wingett, S., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., and Andrews, S. (2015). HiCUP: pipeline for mapping and processing Hi-C data. F1000Research *4*, 1310.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-523.

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics *25*, 1952-1958.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhao, H., Sun, Z., Wang, J., Huang, H., Kocher, J.P., and Wang, L. (2014). CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics *30*, 1006-1007.

Figure 1

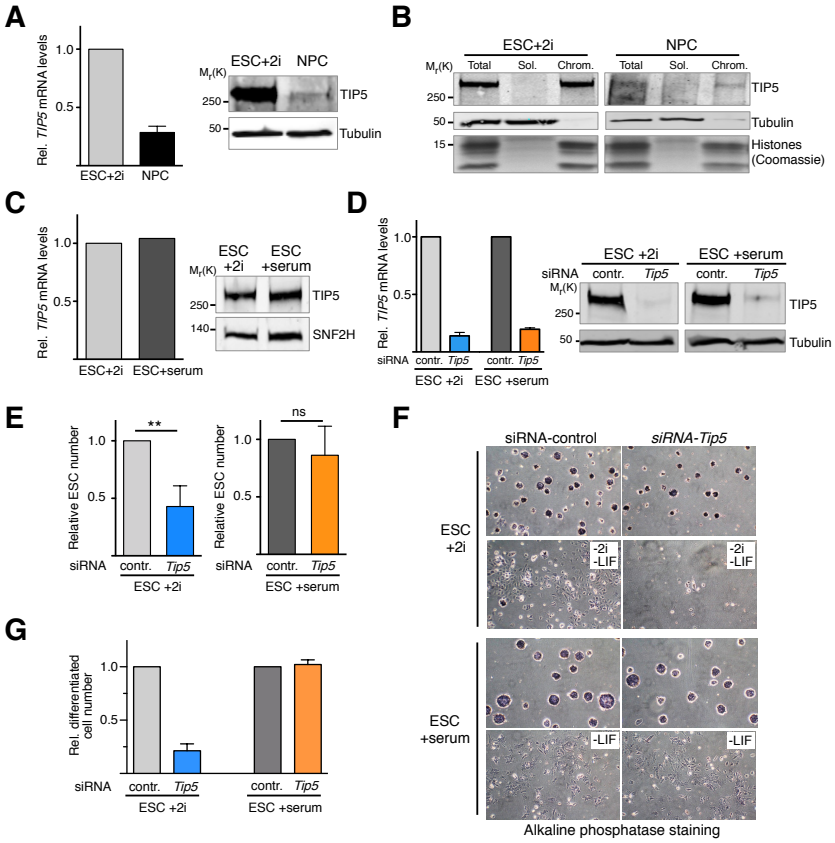

A
Rel. Tif5 mRNA levels
1.0
0.5
ESC+2i
NPC
Mr(K)
ESC+2i
NPC
250
50
TIP5
Tubulin
B
ESC+2i
NPC
Mr(K)
Total
Sol.
Chrom.
Total
Sol.
Chrom.
250
50
15
TIP5
Tubulin
Histones
(Coomassie)
C
Rel. Tif5 mRNA levels
1.0
0.5
ESC+2i
ESC+serum
Mr(K)
ESC
+2i
ESC
+serum
250
140
TIP5
SNF2H
D
Rel. Tif5 mRNA levels
1.0
0.5
siRNA
contr.
Tip5
contr.
Tip5
ESC +2i
ESC +serum
ESC +2i
ESC +serum
siRNA
contr.
Tip5
contr.
Tip5
Mr(K)
250
50
TIP5
Tubulin
E
Relative ESC number
1.0
0.5
**
siRNA
contr.
Tip5
ESC +2i
Relative ESC number
1.0
0.5
ns
siRNA
contr.
Tip5
ESC +serum
F
siRNA-control
siRNA-Tip5
ESC
+2i
-2i
-LIF
-2i
-LIF
ESC
+serum
-LIF
-LIF
Alkaline phosphatase staining
G
Rel. differentiated cell number
1.0
0.5
siRNA
contr.
Tip5
ESC +2i
contr.
Tip5
ESC +serum

Figure 2

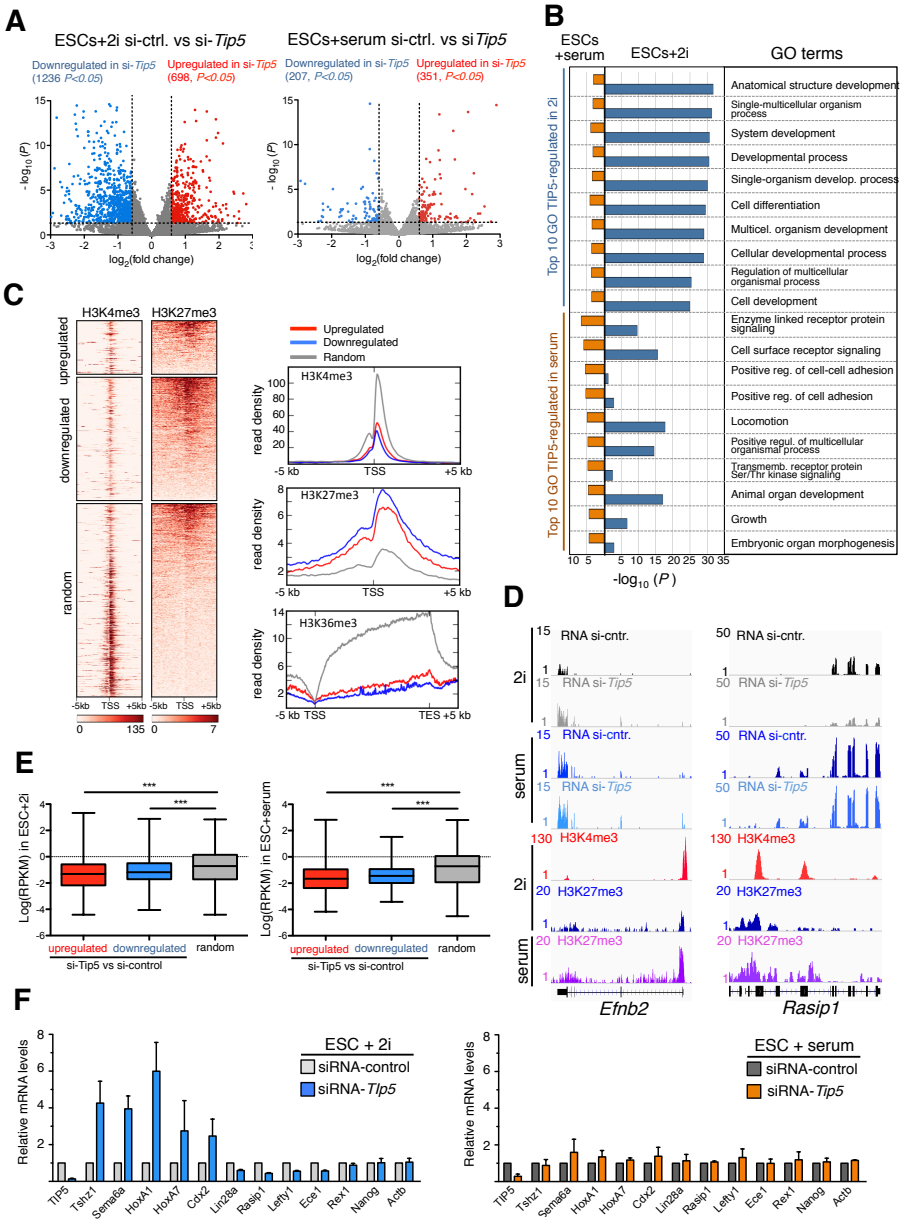

A
ESCs+2i si-ctrl. vs si-Tip5
ESCs+serum si-ctrl. vs si Tip5
Downregulated in si-Tip5 (1236 P<0.05)
Upregulated in si-Tip5 (698, P<0.05)
Downregulated in si-Tip5 (207, P<0.05)
Upregulated in si-Tip5 (351, P<0.05)
- log10 (P)
log2(fold change)
B
ESCs +serum
ESCs+2i
GO terms
Top 10 GO TIP5-regulated in 2i
Top 10 GO TIP5-regulated in serum
Anatomical structure development
Single-multicellular organism process
System development
Developmental process
Single-organism develop. process
Cell differentiation
Multicel. organism development
Cellular developmental process
Regulation of multicellular organismal process
Cell development
Enzyme linked receptor protein signaling
Cell surface receptor signaling
Positive reg. of cell-cell adhesion
Positive reg. of cell adhesion
Locomotion
Positive regul. of multicellular organismal process
Transmemb. receptor protein Ser/Thr kinase signaling
Animal organ development
Growth
Embryonic organ morphogenesis
-log10 (P)
C
H3K4me3
H3K27me3
upregulated
downregulated
random
Upregulated
Downregulated
Random
read density
H3K4me3
H3K27me3
H3K36me3
-5 kb
TSS
+5 kb
TES +5 kb
D
RNA si-cntr.
RNA si-Tip5
2i
serum
H3K4me3
H3K27me3
Efnb2
Rasip1
E
Log(RPKM) in ESC+2i
Log(RPKM) in ESC+serum
upregulated
downregulated
random
si-Tip5 vs si-control
F
ESC + 2i
ESC + serum
siRNA-control
siRNA-Tip5
Relative mRNA levels
Tif5
Tshz1
Sema6a
Hoxa1
Hoxa7
Cdx2
Lin28a
Raslp1
Lefty1
Ece1
Rex1
Nanog
Actb

Figure 3

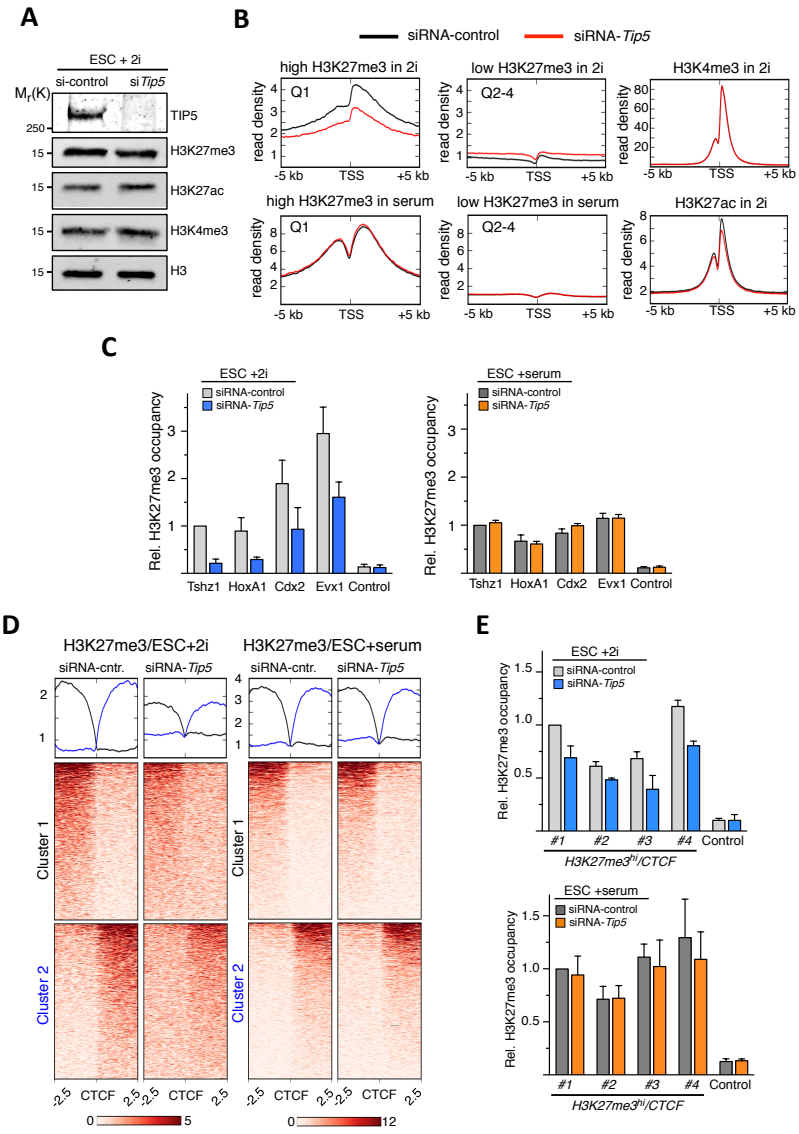

A
ESC + 2i
si-control
si Tip5
Mr(K)
250
TIP5
15
H3K27me3
15
H3K27ac
15
H3K4me3
15
H3
B
siRNA-control
siRNA-Tip5
high H3K27me3 in 2i
low H3K27me3 in 2i
H3K4me3 in 2i
high H3K27me3 in serum
low H3K27me3 in serum
H3K27ac in 2i
Q1
Q2-4
read density
-5 kb
TSS
+5 kb
C
ESC +2i
ESC +serum
siRNA-control
siRNA-Tip5
Rel. H3K27me3 occupancy
Tshz1
HoxA1
Cdx2
Evx1
Control
D
H3K27me3/ESC+2i
H3K27me3/ESC+serum
siRNA-cntr.
siRNA-Tip5
Cluster 1
Cluster 2
CTCF
E
ESC +2i
ESC +serum
siRNA-control
siRNA-Tip5
Rel. H3K27me3 occupancy
#1
#2
#3
#4
Control
H3K27me3hi/CTCF

Figure 4

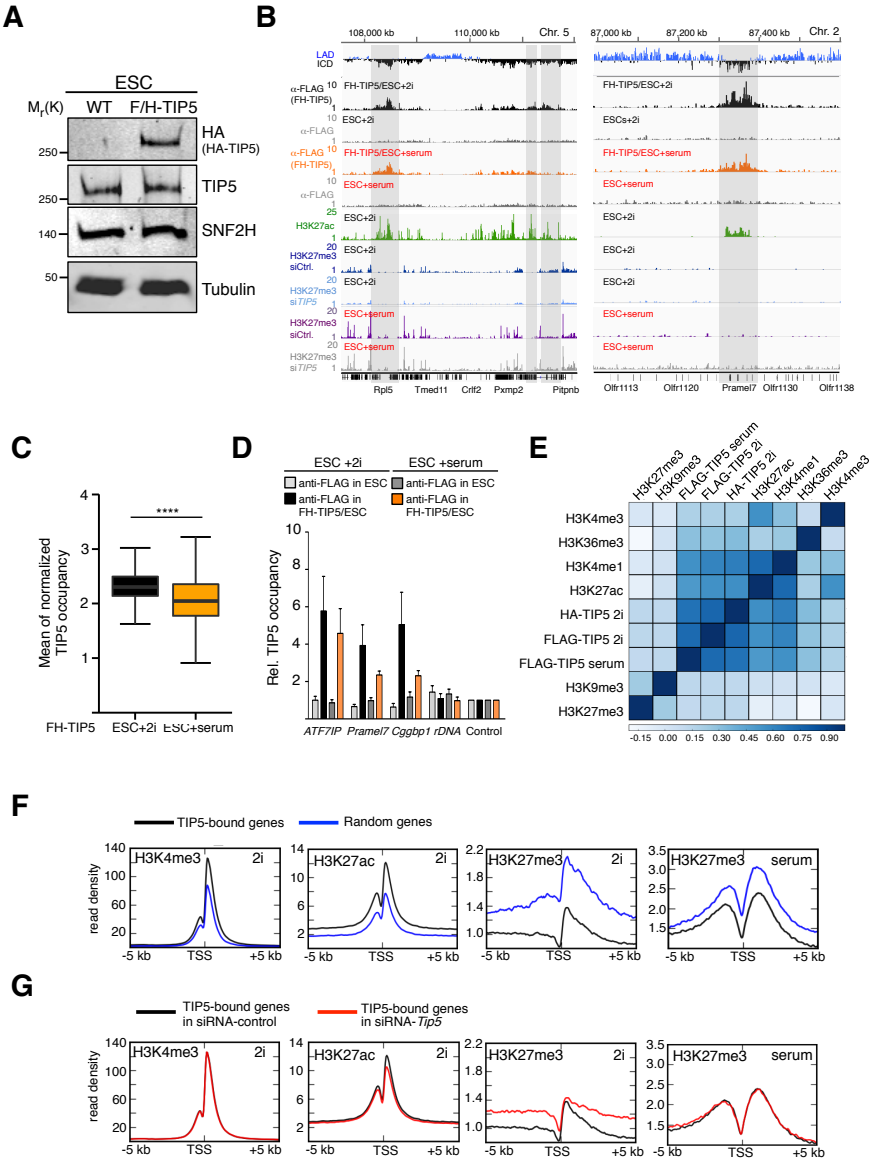

A
ESC
M$_r$(K)
WT
F/H-TIP5
HA
(HA-TIP5)
TIP5
SNF2H
Tubulin
250
250
140
50
B
Chr. 5
Chr. 2
LAD
ICD
α-FLAG
(FH-TIP5)
FH-TIP5/ESC+2i
ESC+2i
FH-TIP5/ESC+serum
ESC+serum
H3K27ac
H3K27me3
siCtrl.
siTIP5
Rpl5
Tmed11
Crlf2
Pxmp2
Pitpnb
Olfr1113
Olfr1120
Pramel7
Olfr1130
Olfr1138
C
Mean of normalized TIP5 occupancy
****
FH-TIP5
ESC+2i
ESC+serum
D
ESC +2i
ESC +serum
anti-FLAG in ESC
anti-FLAG in FH-TIP5/ESC
Rel. TIP5 occupancy
ATF7IP
Pramel7
Cggbp1
rDNA
Control
E
H3K27me3
H3K9me3
FLAG-TIP5 serum
FLAG-TIP5 2i
HA-TIP5 2i
H3K27ac
H3K4me1
H3K36me3
H3K4me3
-0.15 0.00 0.15 0.30 0.45 0.60 0.75 0.90
F
TIP5-bound genes
Random genes
read density
H3K4me3
H3K27ac
H3K27me3
2i
serum
-5 kb
TSS
+5 kb
G
TIP5-bound genes in siRNA-control
TIP5-bound genes in siRNA-Tip5

Figure 5

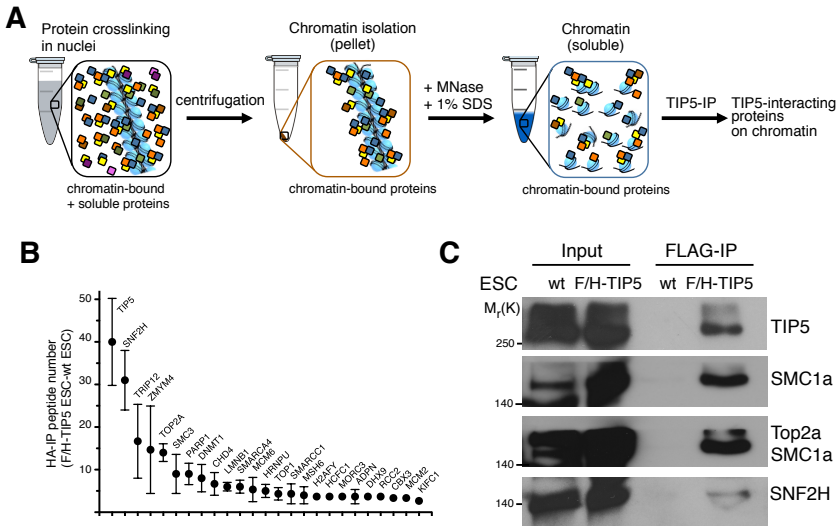

A
Protein crosslinking in nuclei
centrifugation
Chromatin isolation (pellet)
+ MNase
+ 1% SDS
Chromatin (soluble)
TIP5-IP
TIP5-interacting proteins on chromatin
chromatin-bound + soluble proteins
chromatin-bound proteins
chromatin-bound proteins
B
HA-IP peptide number (F/H-TIP5 ESC-wt ESC)
50
40
30
20
10
TIP5
SNF2H
TRIP12
ZMYM4
TOP2A
SMC3
PARP1
DNMT1
CUL4
LMNB1
SMARCA4
MCM6
HNRNPU
TOP1
SMARCC1
MSH6
RZAPLY
RCFC1
MORC3
ADPRN
DHX9
RACC2
CRYX3
MCM2
KIFC1
C
Input
FLAG-IP
ESC
wt
F/H-TIP5
wt
F/H-TIP5
M_r(K)
250
140
140
140
TIP5
SMC1a
Top2a
SMC1a
SNF2H

Figure 6

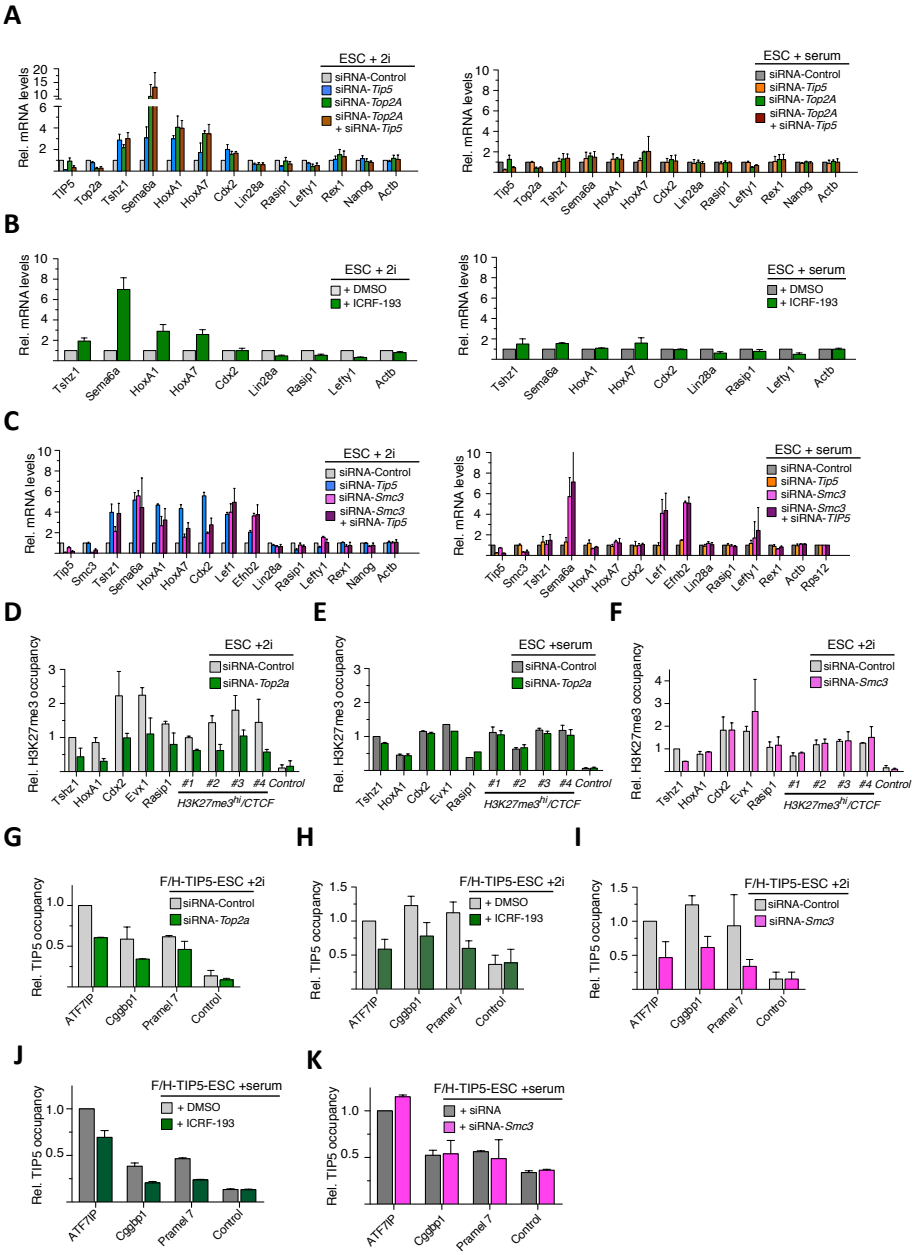
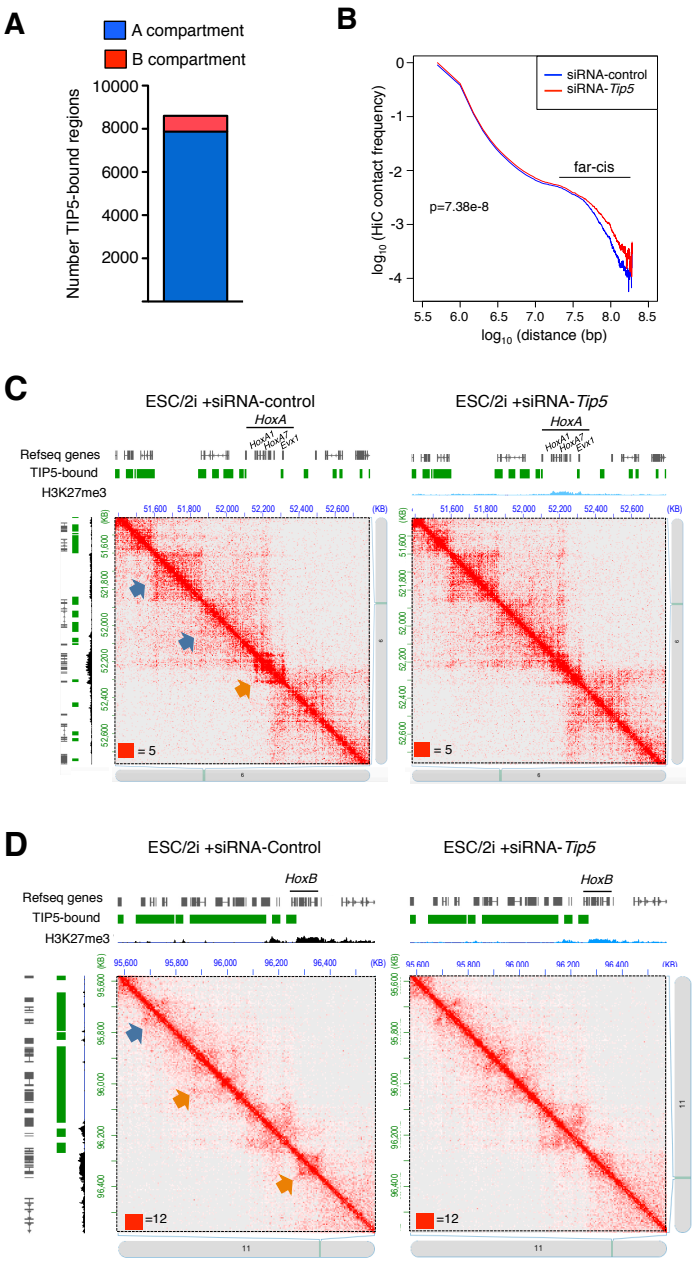
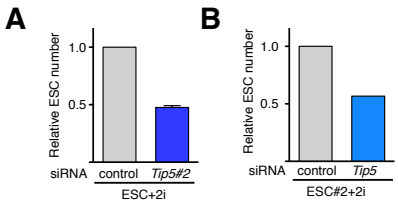

A
ESC + 2i
siRNA-Control
siRNA-Tip5
siRNA-Top2A
siRNA-Top2A + siRNA-Tip5
Rel. mRNA levels
Tip5 Top2a Tshz1 Sema6a HoxA1 HoxA7 Cdx2 Lin28a Rasip1 Lefty1 Rex1 Nanog Actb
ESC + serum
siRNA-Control
siRNA-Tip5
siRNA-Top2A
siRNA-Top2A + siRNA-Tip5
Rel. mRNA levels
Tip5 Top2a Tshz1 Sema6a HoxA1 HoxA7 Cdx2 Lin28a Rasip1 Lefty1 Rex1 Nanog Actb
B
ESC + 2i
+ DMSO
+ ICRF-193
Rel. mRNA levels
Tshz1 Sema6a HoxA1 HoxA7 Cdx2 Lin28a Rasip1 Lefty1 Actb
ESC + serum
+ DMSO
+ ICRF-193
Rel. mRNA levels
Tshz1 Sema6a HoxA1 HoxA7 Cdx2 Lin28a Rasip1 Lefty1 Actb
C
ESC + 2i
siRNA-Control
siRNA-Tip5
siRNA-Smc3
siRNA-Smc3 + siRNA-Tip5
Rel. mRNA levels
Tip5 Smc3 Tshz1 Sema6a HoxA1 HoxA7 Cdx2 Lef1 Efnb2 Lin28a Rasip1 Lefty1 Rex1 Nanog Actb
ESC + serum
siRNA-Control
siRNA-Tip5
siRNA-Smc3
siRNA-Smc3 + siRNA-TIP5
Rel. mRNA levels
Tip5 Smc3 Tshz1 Sema6a HoxA1 HoxA7 Cdx2 Lef1 Efnb2 Lin28a Rasip1 Lefty1 Rex1 Actb Rps12
D
ESC +2i
siRNA-Control
siRNA-Top2a
Rel. H3K27me3 occupancy
Tshz1 HoxA1 Cdx2 Evx1 Rasip1 #1 #2 #3 #4 Control
H3K27me3hi/CTCF
E
ESC +serum
siRNA-Control
siRNA-Top2a
Rel. H3K27me3 occupancy
Tshz1 HoxA1 Cdx2 Evx1 Rasip1 #1 #2 #3 #4 Control
H3K27me3hi/CTCF
F
ESC +2i
siRNA-Control
siRNA-Smc3
Rel. H3K27me3 occupancy
Tshz1 HoxA1 Cdx2 Evx1 Rasip1 #1 #2 #3 #4 Control
H3K27me3hi/CTCF
G
F/H-TIP5-ESC +2i
siRNA-Control
siRNA-Top2a
Rel. TIP5 occupancy
ATF7IP Cggbp1 Pramel 7 Control
H
F/H-TIP5-ESC +2i
+ DMSO
+ ICRF-193
Rel. TIP5 occupancy
ATF7IP Cggbp1 Pramel 7 Control
I
F/H-TIP5-ESC +2i
siRNA-Control
siRNA-Smc3
Rel. TIP5 occupancy
ATF7IP Cggbp1 Pramel 7 Control
J
F/H-TIP5-ESC +serum
+ DMSO
+ ICRF-193
Rel. TIP5 occupancy
ATF7IP Cggbp1 Pramel 7 Control
K
F/H-TIP5-ESC +serum
+ siRNA
+ siRNA-Smc3
Rel. TIP5 occupancy
ATF7IP Cggbp1 Pramel 7 Control

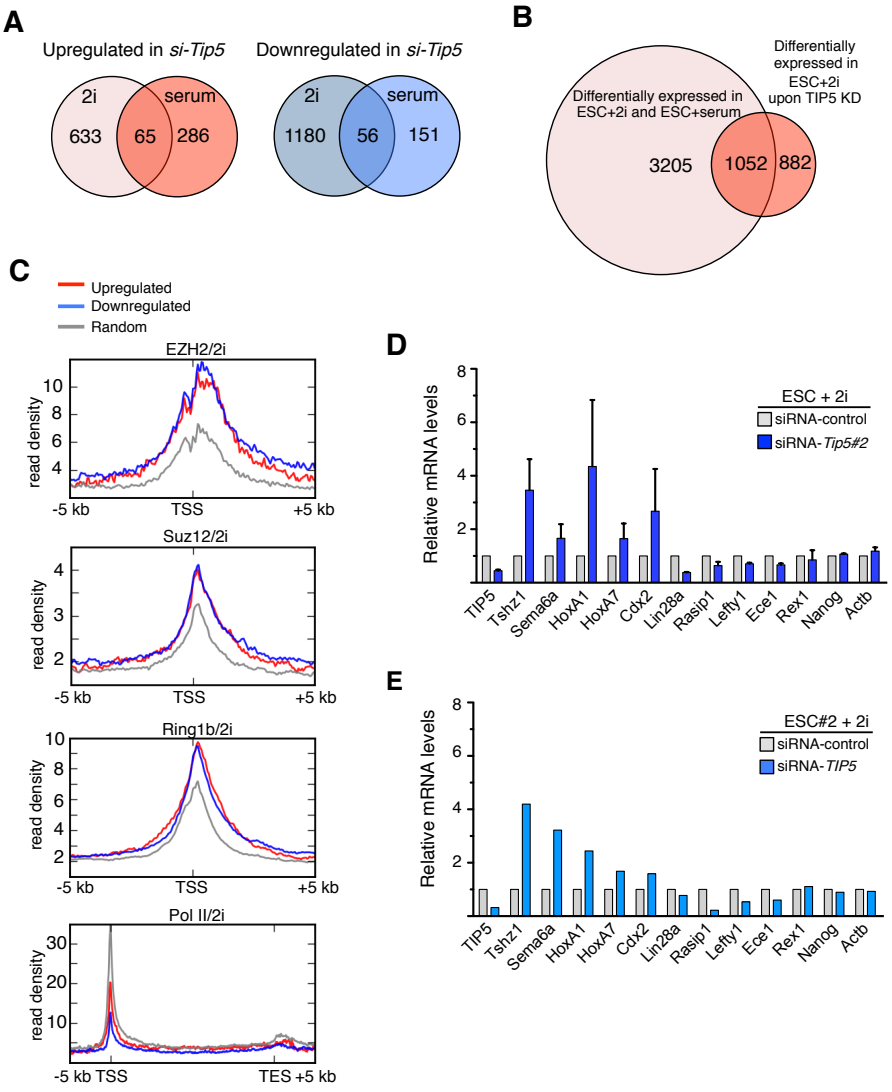
Figure 7


A
A compartment
B compartment
Number TIP5-bound regions
10000
8000
6000
4000
2000
B
siRNA-control
siRNA-Tip5
far-cis
p=7.38e-8
log10 (HiC contact frequency)
log10 (distance (bp)
C
ESC/2i +siRNA-control
ESC/2i +siRNA-Tip5
HoxA
Refseq genes
TIP5-bound
H3K27me3
D
ESC/2i +siRNA-Control
ESC/2i +siRNA-Tip5
HoxB

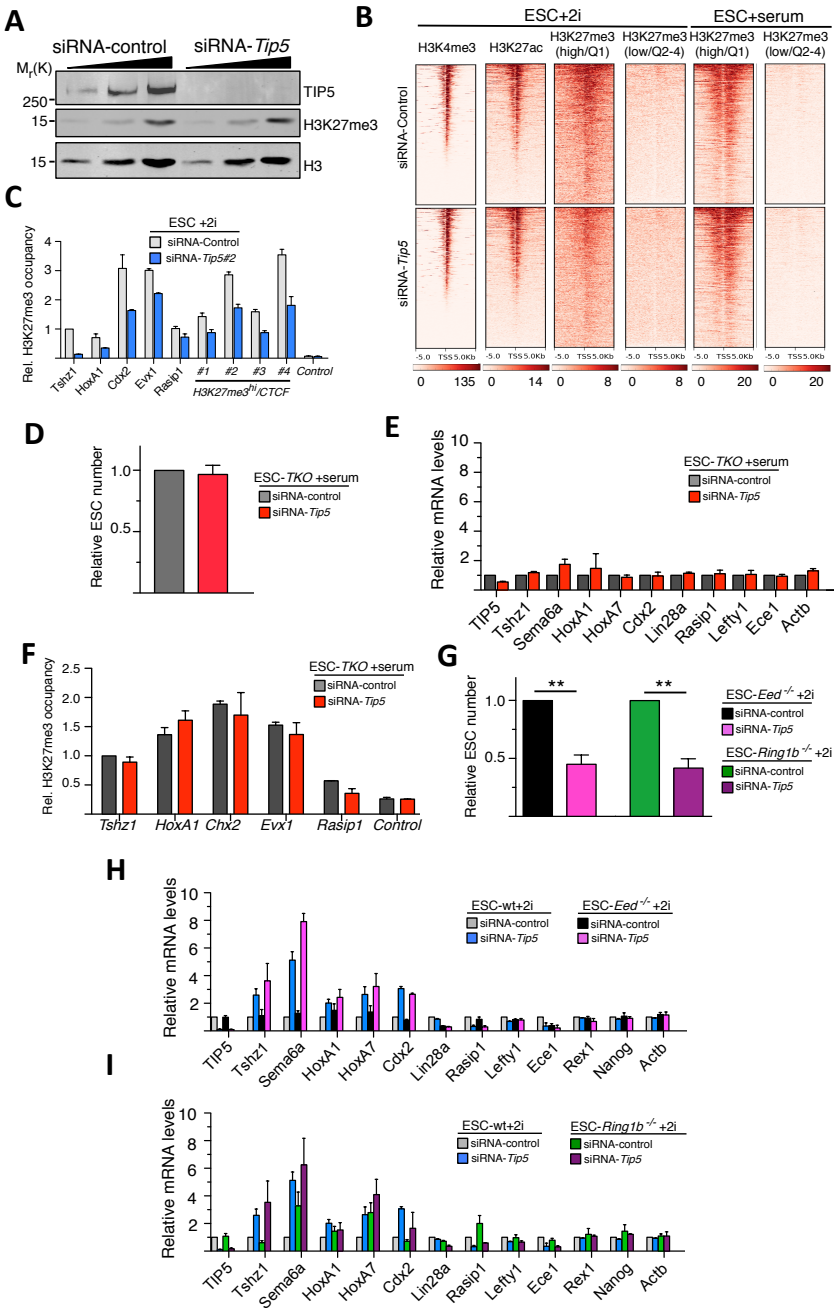
Supplementary Figure 1



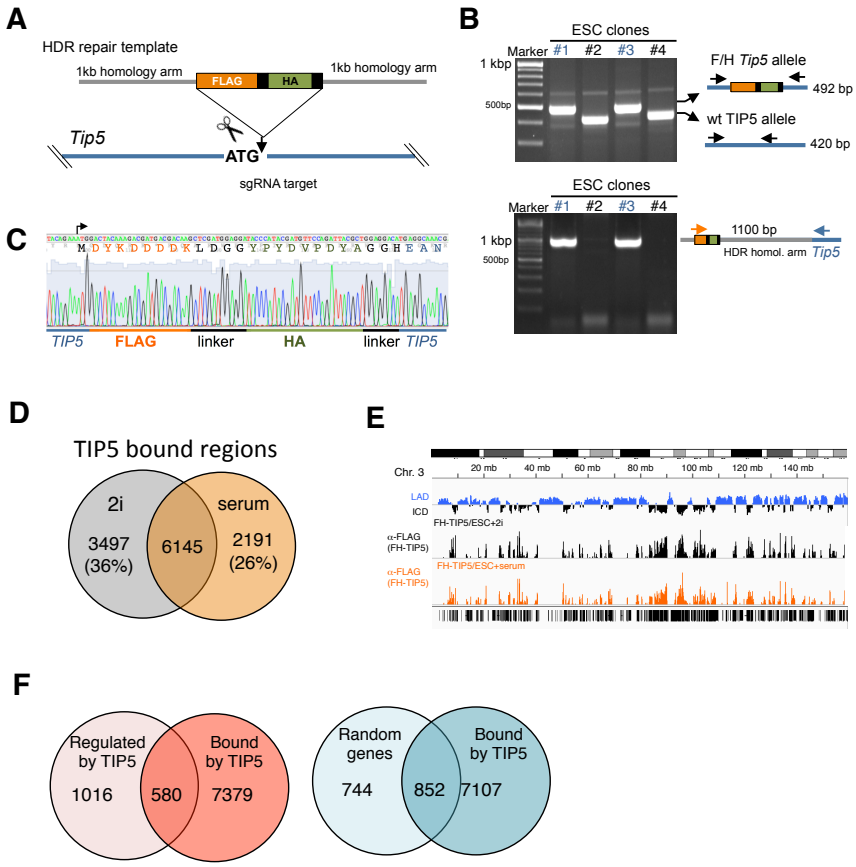
Supplementary Figure 2



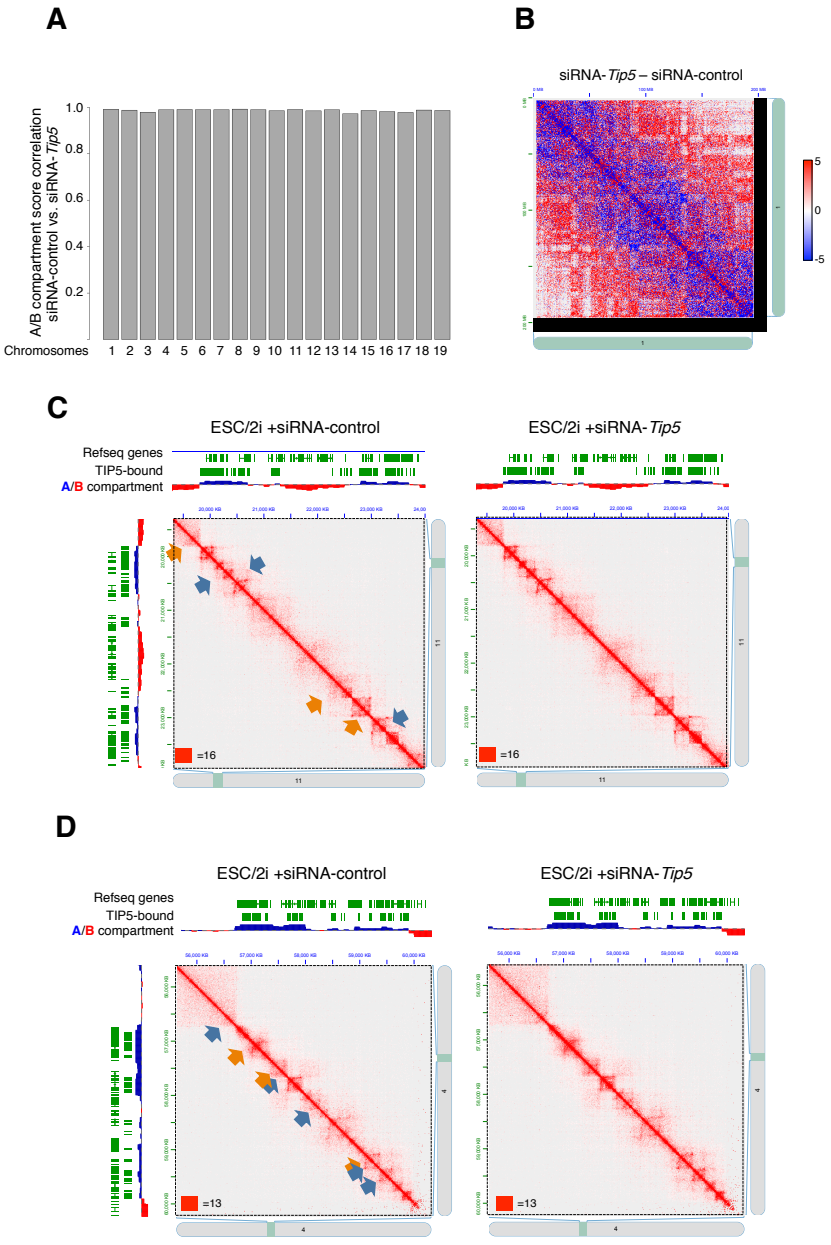
Supplementary Figure 3



Supplementary Figure 4



Supplementary Figure 5



Published online: June 6, 2017

Article

SOURCE DATA | TRANSPARENT PROCESS

EMBO reports

The RNA helicase DHX9 establishes nucleolar heterochromatin, and this activity is required for embryonic stem cell differentiation

Sergio Leone[1,2], Dominik Bär[1], Coenraad Frederik Slabber[1], Damian Dalcher[1,2] & Raffaella Santoro[1,*]

Abstract

Long non-coding RNAs (lncRNAs) have been implicated in the regulation of chromatin conformation and epigenetic patterns. lncRNA expression levels are widely taken as an indicator for functional properties. However, the role of RNA processing in modulating distinct features of the same lncRNA is less understood. The establishment of heterochromatin at rRNA genes depends on the processing of IGS-rRNA into pRNA, a reaction that is impaired in embryonic stem cells (ESCs) and activated only upon differentiation. The production of mature pRNA is essential since it guides the repressor TIP5 to rRNA genes, and IGS-rRNA abolishes this process. Through screening for IGS-rRNA-binding proteins, we here identify the RNA helicase DHX9 as a regulator of pRNA processing. DHX9 binds to rRNA genes only upon ESC differentiation and its activity guides TIP5 to rRNA genes and establishes heterochromatin. Remarkably, ESCs depleted of DHX9 are unable to differentiate and this phenotype is reverted by the addition of pRNA, whereas providing IGS-rRNA and pRNA mutants deficient for TIP5 binding are not sufficient. Our results reveal insights into lncRNA biogenesis during development and support a model in which the state of rRNA gene chromatin is part of the regulatory network that controls exit from pluripotency and initiation of differentiation pathways.

Keywords DHX9; embryonic stem cells; heterochromatin; lncRNA
Subject Categories Chromatin, Epigenetics, Genomics & Functional Genomics; RNA Biology; Stem Cells
DOI 10.15252/embr.201744330 | Received 5 April 2017 | Revised 22 April 2017 | Accepted 25 April 2017 | Published online 6 June 2017
EMBO Reports (2017) 18: 1248–1262

Introduction

The nucleolus is the compartment where ribosome biogenesis takes place, a process that is initiated by transcription of ribosomal RNA (rRNA) genes that synthesize ribosomal rRNA [1]. In each somatic cell, a fraction of the 400 rRNA genes is transcriptionally silent and organized in heterochromatic structures, including CpG methylation and histone H3K9 methylation [2,3]. However, until now the role of silent heterochromatic rRNA repeats is not yet understood. Indeed, the formation of heterochromatin at rRNA genes appears to not be implicated in ribosome biogenesis since silent rRNA repeats do not become transcriptionally activated even under conditions of high metabolic activities of the cell [4].

The higher order organization of genomes is functionally important to ensure correct execution of gene expression programs. For instance, as cells differentiate into specialized cell types, chromosomes undergo diverse structural and organizational changes that affect gene expression and other cellular functions [5]. The chromatin of embryonic stem cells (ESCs) is largely devoid of compact heterochromatin blocks when compared to lineage-committed cells [6,7]. This structure well reflects the plasticity and transcriptional permissiveness of ESC genome that need to maintain the ability to enter any distinct transcriptional programs for lineage specification [8,9]. Upon differentiation, a large portion of the ESC genome remodels toward a highly condensed heterochromatic form [10]. These changes are also accompanied by alterations of nuclear architecture such as the clustering of centromeres either at the nucleolus or at the nuclear envelope [11,12]. However, it still remains elusive how the switch from a lower to a higher order chromatin structure is achieved during ESC differentiation and whether this process plays a role in ESC differentiation.

Previous results have shown that also rRNA genes undergo chromatin changes during mouse ESC differentiation [13,14]. In ESCs, all 400 rRNA gene copies are active and only upon differentiation a fraction of these genes acquire heterochromatic marks such as CpG methylation and histone H3K9 methylation, as found in somatic cells [2,13]. In somatic mouse and human cells, establishment of heterochromatin at rRNA genes is mediated by TIP5 through its association with the long non-coding (lnc)RNA pRNA, DNA methyltransferases, and histone modifier enzymes [3,15–17]. Although mouse and human pRNA sequences are lowly conserved in their sequences, they display a similar structure that serves for the association with TIP5 *in vitro* and *in vivo* [15,16,18]. pRNA derives from the processing of the 2-kb-long IGS-rRNA, a transcript originating from the spacer promoter that, in mouse, is located 2-kb upstream

1 Department of Molecular Mechanisms of Disease, University of Zurich, Zurich, Switzerland
2 Molecular Life Science Program, Life Science Zurich Graduate School, University of Zurich, Zurich, Switzerland
*Corresponding author. Tel: +41 44 63 55475; E-mail: raffaella.santoro@dmmd.uzh.ch

Published online: June 6, 2017

the main rRNA gene promoter [16,19]. The production of mature pRNA is essential for the formation of heterochromatin at rRNA genes since only mature pRNA can guide TIP5 to nucleolus to establish silencing while the association of TIP5 with the unprocessed transcript abolishes this process [13]. Mature pRNA promotes the association of TIP5 with TTF1, a nucleolar docking factor bound to the main promoter of rRNA genes, whereas IGS-rRNA destroys the interaction with TTF1 and impairs TIP5 recruitment [13,20]. In ESCs, the maturation of IGS-rRNA into pRNA is abolished and it is only upon ESC differentiation that the IGS-rRNA processing is activated to produce pRNA, which leads to the establishment heterochromatin at a fraction of rRNA genes for the first time. Remarkably, addition of mature pRNA in ESCs was not only sufficient to establish nucleolar heterochromatin and to downregulate rRNA synthesis but also primed ESCs for differentiation. These results are consistent with recent studies proposing that elevated rRNA biosynthesis sustains pluripotency in mouse and human ESCs since rRNA downregulation through chemical inhibition of Pol I activity or deletion of fibrillarin induces differentiation [21,22].

The critical role of pRNA biogenesis in the establishment of rRNA gene heterochromatin represents an important example of how the different features of the same lncRNA can be modulated to regulate chromatin conformation and epigenetic patterns during development. However, how this process is regulated is not understood. To this end, we performed a screening for IGS-rRNA-binding proteins and identified the RNA helicase DHX9 to be required for the production of mature pRNA. DHX9 binds to rRNA genes only upon ESC differentiation, and its activity to process IGS-rRNA into pRNA is necessary to guide TIP5 to rRNA genes and establish rRNA gene heterochromatin. Remarkably, ESCs depleted of DHX9 are unable to differentiate and this phenotype can be reverted by the sole addition of mature pRNA. Taken together, these results provide molecular insights into lncRNA biogenesis that modulates different features of the same lncRNA and highlights the importance of lncRNA processing for the establishment of chromatin states during development.

Results

The RNA helicase DHX9 associates with IGS-rRNA

pRNA is a 250–300 nucleotide transcript that derives from processing of the 2-kb-long IGS-rRNA [16,19]. Production of pRNA is a key step for the formation of heterochromatin at rRNA genes since only mature pRNA can guide TIP5 to rRNA genes to establish repressive chromatin states [13]. We sought to identify factors implicated in IGS-rRNA processing by searching for proteins able to specifically bind to IGS-rRNA. We performed GRNA chromatography, a method that relies on the interaction between a 19 nt RNA element called *BoxB* and a 22 amino acid long peptide λ_{N22} [23]. We converted Glutathione Sepharose into an RNA affinity matrix by binding GST-λ_{N22} fusion protein to *in vitro* synthesized *BoxB*-IGS-rRNA or *BoxB-Control-RNA* and performed pull-downs using nuclear extracts of NIH 3T3 or neural progenitors cells (NPCs), both proficient for IGS-rRNA processing [13,19] (Fig 1A). Proteins bound to IGS-rRNA and control-RNA sequences were eluted with RNase A and analyzed by mass spectrometry (Figs 1B and EV1A–C, Dataset EV1). To identify factors that specifically associate with IGS-rRNA sequences, we considered only proteins that were present in all the experiments with a minimum of two peptides and that were associated with at least twofold higher peptide content to IGS-rRNA compared to control-RNA (Figs 1B and EV1B and C). Only six proteins fulfilled these criteria: DHX9, FUBP1, FUBP3, CSTF1, TIAR, and LSM6. As expected, all these factors are related to RNA pathways. The RNA helicase DHX9 (RNA Helicase A, RHA) caught our attention due to its reported localization in nucleoli of mouse and human cells, the compartment where processing of IGS-rRNA takes place [24–26]. Accordingly, immunofluorescence analysis revealed the presence of endogenous DHX9 or ectopic GFP-DHX9 within nucleoli, as evidenced by the co-localization with the nucleolar upstream binding factor UBF (Fig 1C). Moreover, in our effort to identify factors associated with TIP5 through immuno-precipitation (IP) combined with mass-spec analysis, DHX9 was consistently detected as a TIP5-interacting protein (data not shown). Accordingly, co-immunoprecipitation of TIP5 in HEK293T cells revealed the association of TIP5 with endogenous DHX9 and the N-terminus domain of TIP5 was sufficient for this interaction (Fig 1D). The interaction between endogenous TIP5 and DHX9 was also detected in ESCs (Fig EV1D). Finally, the identification of DHX9-interacting proteins in ESCs, ESCs after 4 days of differentiation, NIH 3T3 and HEK293 cells through anti-FLAG immuno-precipitation of ectopically expressed FLAG-DHX9 followed by mass spectrometry revealed a strong enrichment in factors implicated in RNA processing (Datasets EV2 and EV3, and Fig EV2). In particular, components of the spliceosome (as determined by STRING database of interaction and KEGG pathway database) were remarkably abundant [27, 28]. Consistent with these results, spliceosome was the major hit obtained through cellular component and pathway analysis in each of the GRNA chromatography experiments performed to identify IGS-rRNA-binding proteins (Dataset EV1). Interestingly, the absence of the spliceosome canonical DExH/RHA helicases (i.e., PRPF 2, 16, 22, and 43) [29] among DHX9 and IGS-rRNA interacting factors suggests that DHX9 does not interact with the canonical spliceosome complex.

DHX9 belongs to the DExH/RHA family of helicase superfamily 2 and is characterized by two copies of a double-stranded RNA-binding domain (DSRM) at the amino terminus, a helicase core domain (HrpA) in the central region, and an RGG-rich region at the carboxyl terminus, which confers both RNA and DNA helicase activities [30–32]. DHX9 is the homolog of the maleless gene product MLE in *Drosophila*, where it is essential for dosage compensation between the two X chromosomes of females and the single one of males [33]. In mammalian cells, DHX9 has been implicated in a variety of processes such as transcriptional activation [34–38] and genome stability [39,40]. However, the co-activating function of DHX9 does not necessarily require ATP-hydrolysis or DNA unwinding [36,38].

The ability of DHX9 to bind to IGS-rRNA, its nucleolar localization, the association with TIP5 and its link to factors implicated in spliceosome formation prompted us to investigate whether DHX9 plays a role in the processing of IGS-rRNA. We initially employed a previously established assay based on the transfection of the IGS-rRNA reporter plasmid in NIH 3T3 cells, proficient for IGS-rRNA processing (Fig 2A) [13]. Transcripts are measured by strand-specific reverse transcription (RT) using a primer hybridizing vector sequences downstream the main rRNA gene promoter followed by

Published online: June 6, 2017
EMBO reports

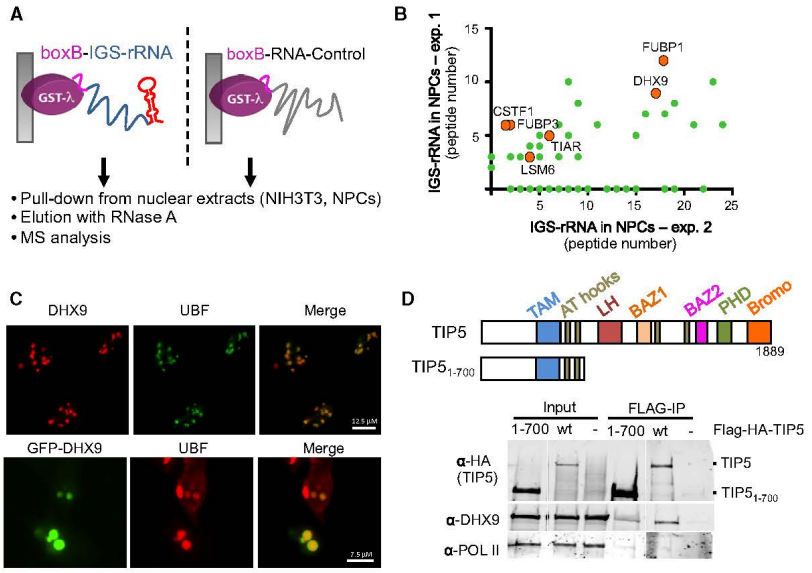


Figure 1. DHX9 is a nucleolar protein that associates with IGS-rRNA and TIP5.

A Schema represents the GRNA chromatography employed for the identification of IGS-rRNA-binding proteins.

B Representative mass spectrometry analysis of proteins pulled down with IGS-rRNA and found enriched in all three independent experiments (labeled with orange circles), using nuclear extracts from NIH 3T3 cells or NPCs. Values (green circles) refer to peptide numbers obtained in each experiment. For direct comparison with the other experiments, the same data are shown in Fig EV1B.

C DHX9 localizes in nucleoli. Immunofluorescence analysis with DHX9 antibodies in U2OS cells and with GFP antibodies in NIH 3T3 cells transfected with GFP-DHX9 plasmids. Nucleoli are visualized with anti-UBF.

D Endogenous DHX9 associates with TIP5. Schema represents the domain organization of TIP5. Anti-FLAG immunoprecipitation from HEK293T cells expressing FLAG-HA-TIP5 or FLAG-HA-TIP5$_{1-700}$. Immunoblots show association of TIP5 with DHX9 but not with the RNA Pol II.

Source data are available online for this figure.

amplification of pRNA and 5′-IGS-rRNA regions. Since IGS-rRNA also includes pRNA sequences, values obtained through amplification with primers amplifying pRNA sequences represent the total amount of mature pRNA and unprocessed IGS-rRNA transcripts whereas measurements of 5′-IGS-rRNA sequences correspond to the levels of unprocessed transcripts only. We made use of the IGS-rRNA reporter assay in NIH 3T3 cells depleted of DHX9 by siRNA and found that processing of ectopic IGS-rRNA was less efficient when compared to control cells (Fig 2A). Knockdown of DHX9 also induced a consistent accumulation of endogenous unprocessed IGS-rRNA (Fig 2B). Depletion of DHX9 did not affect the total levels of IGS-rRNA and pRNA sequences, indicating that DHX9 is not implicated in the synthesis of IGS-rRNA itself but most likely acts on its processing (Fig EV1E and F). Remarkably, the defects in IGS-rRNA processing upon DHX9 knockdown could be reversed upon ectopic expression of GFP-DHX9, which is not targeted by the siRNA-DHX9 (Fig 2C). Expression of GFP-DHX9$_{K417R}$ that contains a mutation impairing the ATP-helicase activity [36,41] displayed a similar efficiency for restoring IGS-rRNA processing (Fig 2C), suggesting that DHX9 RNA helicase activity is not required for IGS-rRNA maturation. Together, these results indicate that DHX9 is implicated in the maturation of IGS-rRNA into pRNA.

IGS-rRNA is composed of three sequence elements, namely spacer region, enhancer repeats and pRNA, [16,19] (Fig 3A). To determine whether DHX9 has a preferential association with one of these sequences, we performed electrophoretic mobility shift (EMSA) competition assays. Increasing amounts of *in vitro* transcripts corresponding to selected regions of murine rRNA were used to compete for binding of recombinant TIP5 or DHX9 to radiolabeled runoff transcripts from pBluescript (RNA$_{MCS}$) [18]. Consistent with previous results, TIP5 displayed a higher binding affinity for pRNA than a control-RNA sequence since only pRNA can efficiently compete for the binding to TIP5 (Fig 3B) [13,16,18]. DHX9 bound to RNA, forming high-molecular-weight complexes. However, the specificity of DHX9 for pRNA was lower than the one observed for TIP5. Indeed, control-RNA could also compete for binding, although to a less extent since higher amounts of transcripts were required (Fig 3C). The interaction of DHX9 with nucleic acids is specific for

Published online: June 6, 2017

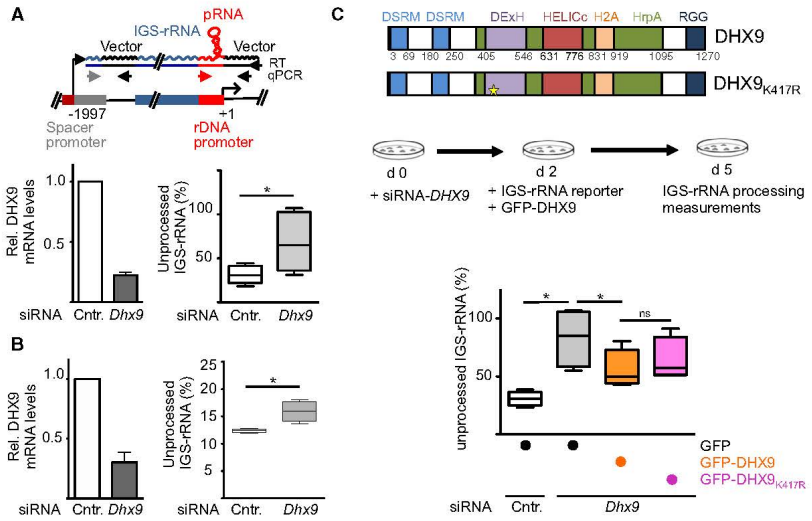


Figure 2. DHX9 is implicated in IGS-rRNA processing.

A Schema depicts the IGS-rRNA reporter plasmid. Black arrows represent primers used to perform RT or to amplify plasmid sequences. Gray and red arrows indicate primers hybridizing to rRNA sequences. siRNA-DHX9 or siRNA-control-treated NIH 3T3 cells were transfected with IGS-rRNA reporter plasmid. Transcripts are measured by strand-specific reverse transcription (RT) using a primer hybridizing vector sequences downstream the main rRNA gene promoter followed by amplification of pRNA (red and black arrows) and 5′-IGS-rRNA regions (gray and black arrows). Data from six experiments are represented as values of amplifications of the 5′-region of IGS-rRNA (unprocessed) normalized to amplifications of pRNA sequences (unprocessed + processed) of the reporter plasmid. Box plots depict the minimum and maximum values. The median is represented by a horizontal line within the boxes. *DHX9* knockdown efficiency is shown on the left. Transcripts were normalized to *GAPDH* mRNA levels and siRNA-control samples (mean ± SD).

B Measurements of endogenous IGS-rRNA levels in NIH 3T3 cells depleted of DHX9 by siRNA. Data from four experiments have been measured for endogenous transcripts as described in (A).

C Schema depicts the domain organization of DHX9 and the strategy used to measure processing upon expression of plasmids expressing GFP-DHX9 in NIH 3T3 cells depleted of DHX9 by siRNA. Values are from four independent experiments and were calculated as described in (A).

Data information: Statistical significance (*P*-values) for the experiments was calculated using the paired two-tailed *t*-test (*$P < 0.05$; ns, non-significant).

RNA as shown by the higher affinity for pRNA than rDNA promoter sequences (Fig 3D). DHX9 displayed a slight preferential binding for spacer and enhancer RNA sequences as compared to pRNA (Fig 3E). These results are consistent with RNA immunoprecipitation (RIP) analyses of formaldehyde cross-linked NIH 3T3 cells showing that DHX9 directly associates with IGS-rRNA-specific sequences but not with pRNA and that this interaction depends on its RNA-binding domains DSRM and RGG (Fig 3F and G). Together these data support the GRNA chromatography results identifying DHX9 as an IGS-rRNA-specific binding factor. Finally, the preferential association of DHX9 for IGS-rRNA sequences upstream of the pRNA element is further supported by the lack of DHX9 in GRNA chromatography experiments using pRNA as bait (data not shown).

DHX9 associates with TIP5 independently of RNA

The ability of DHX9 to bind IGS-rRNA prompted us to investigate whether RNA mediates the association of DHX9 with TIP5. We performed co-immunoprecipitations from HEK293T cells transfected with plasmids expressing Flag-HA-TIP5 and GFP-DHX9 or GFP-DHX9 mutants ($DHX9_{K417R}$, $DHX9_{\Delta DSRM}$, $DHX9_{\Delta RGG}$; Fig 4A). As shown in Fig 4B, DHX9 mutations inactivating the ATPase activity ($GFP\text{-}DHX9_{K417R}$) did not affect the association with TIP5. Remarkably, RNase A treatment did not alter DHX9-TIP5 interaction. Consistent with these results, the RNA-binding domains of DHX9 (DSRM and RGG) were not required for TIP5-DHX9 interaction (Fig 4C). Although co-expression of $DHX9_{\Delta DSRM}$ and $DHX9_{\Delta RGG}$ mutants negatively affects expression of ectopic Flag-HA-TIP5 (but not of the endogenous protein, Fig EV3), their association with TIP5 was similar to the one observed with wild-type DHX9. Similarly, $TIP5_{W531G, Y532A}$, a mutant with impaired RNA-binding ability [16], efficiently interacted with DHX9 (Fig 4D). Taken together, these results indicate that the association of DHX9 with TIP5 does not require RNA. Considering the preferential binding of TIP5 to pRNA sequences [13] and the association of DHX9 with IGS-rRNA, it may be possible the TIP5-DHX9 complex can simultaneously associate with IGS-rRNA through the interaction of TIP5 with pRNA sequences and the binding of DHX9 to the rest of the unprocessed transcript.

Published online: June 6, 2017
EMBO reports

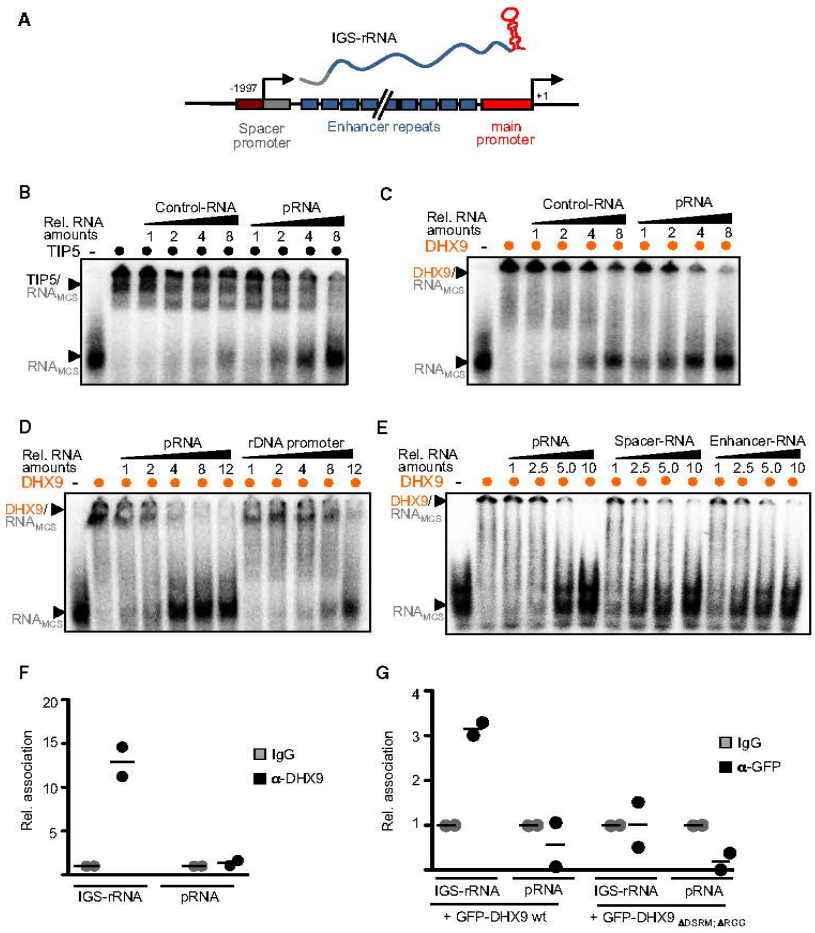


Figure 3. DHX9 associates with RNA.

A Schema depicts the mouse 5′-rRNA gene organization and IGS-rRNA sequence composition: Spacer promoter (gray), enhancer repeats (blue), main promoter (red, pRNA), and transcription start sites of IGS-rRNA (−1,997) and 45S pre-rRNA (+1).

B Increasing equal moles of *in vitro* transcripts corresponding to control-RNA or pRNA (starting from 0.32 pmoles) were used to compete for binding of 50 ng recombinant $TIP5_{1-600}$ to radiolabeled runoff transcripts from pBluescript (RNA_{MCS}). RNA/protein complexes were analyzed by EMSA.

C Binding 75 ng of recombinant DHX9 to RNA was analyzed as described in (B). Competition was performed using increasing equal moles of *in vitro* transcripts corresponding to control-RNA or pRNA (starting from 0.08 pmoles).

D DHX9 preferentially binds to RNA than DNA. The same nucleotide sequences as RNA (pRNA) or double-stranded DNA (rDNA; starting from 0.08 pmoles) were used to compete for binding of 75 ng DHX9 to radiolabeled MCS-RNA.

E The binding of 75 ng DHX9 was analyzed with equal moles of pRNA, spacer, and enhancer repeat transcripts (starting from 0.08 pmoles).

F DHX9 associates *in vivo* with IGS-rRNA. RNA immunoprecipitation (RIP) analysis from formaldehyde cross-linked NIH 3T3 cells. Scatter plot represents the values and the mean of two independent experiments.

G DHX9 RNA-binding domains are required for the association with IGS-rRNA. RIP analysis from NIH 3T3 cells transfected with GFP-DHX9 and GFP-$DHX9_{\Delta DSRM-\Delta RGG}$ expression plasmids. RNA was measured by qPCR through amplification of IGS-rRNA (−1,997/−1,905) and pRNA (−165/−1) sequences. Values are normalized to input and IgG values from two independent experiments. Scatter plot represents the values and the mean of two independent experiments.

Published online: June 6, 2017

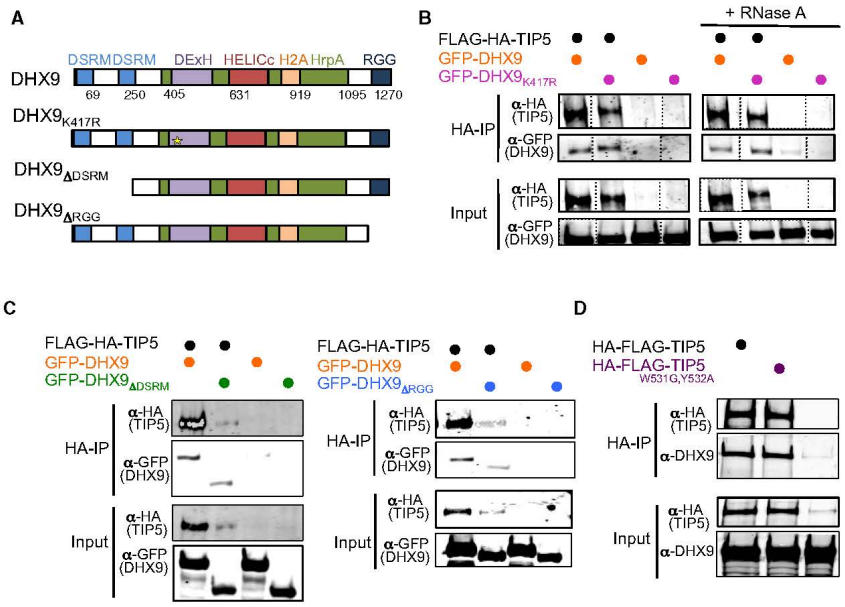


Figure 4. TIP5-DHX9 interaction is not mediated by RNA.

A Schema represents the domain organization of DHX9 and the mutants used for the interaction study.

B Anti-FLAG immunoprecipitation from HEK293T cells expressing FLAG-HA-TIP5 and GFP-DHX9 or GFP $DHX9_{K417R}$. Bead-bound FLAG-TIP5 immunoprecipitates were incubated with RNase A. After washing, co-precipitated proteins were visualized with anti-HA and -GFP antibodies. The data show one representative experiment out of two independent experiments.

C Anti-FLAG immunoprecipitation from HEK293T cells expressing FLAG-HA-TIP5, GFP-DHX9, and GFP-DHX9 mutants ΔDSRM (left panel) or ΔRGG (right panel).

D Anti-FLAG immunoprecipitation from HEK293T cells expressing FLAG-HA-TIP5 or the RNA-binding deficient mutant FLAG-HA-$TIP5_{W513G,Y532A}$.

Source data are available online for this figure.

DHX9 is required for the formation of heterochromatin at rRNA genes through processing of IGS-rRNA into pRNA

The results above indicated that DHX9 is required for processing of IGS-rRNA into pRNA. Previous findings showed that only mature pRNA could guide TIP5 to rRNA genes by promoting the association of TIP5 with TTF1, a nucleolar docking factor bound to the main promoter of rRNA genes [13]. The unprocessed IGS-rRNA alone, on the other hand, destroys the interaction with TTF1 and impairs TIP5 recruitment to rRNA genes. We therefore investigated whether depletion of DHX9 affects the association of TIP5 with rRNA genes, TIP5 nucleolar localization, and rRNA gene silencing. We analyzed NIH 3T3 cells, in which the function of TIP5 and pRNA in establishing silencing at 40–50% of rRNA genes has been well characterized [3,16,19]. While knockdown of DHX9 did not alter TIP5 levels (Fig 5A), TIP5 occupancy at the rRNA genes was impaired in the absence of DHX9 (Fig 5B). Consistent with these results, the nucleolar localization of TIP5 was lost upon depletion of DHX9 (Fig 5C). Remarkably, upon transfection of pRNA in cells depleted of DHX9 the localization of TIP5 was retained in nucleoli (Fig 5D). Taken together, these results indicate that DHX9 is required for the association of TIP5 with rRNA genes through its activity in processing IGS-rRNA into pRNA. We then analyzed whether the nucleolar localization of DHX9 depends on TIP5. Since all DHX9 antibodies tested so far failed to detect a specific signal of DHX9 in mouse cells by immunofluorescence analyses (data not shown), we measured the cellular localization of DHX9 upon TIP5 knockdown in U2OS cells (human osteosarcoma cells). U2OS cells showed a strong enrichment and specific signal of DHX9 in nucleoli (Figs 1C and 5E). Knockdown of TIP5 did not affect DHX9 levels (Fig EV4); however, the enrichment of DHX9 in nucleoli was lost and DHX9 was mostly localized in nucleoplasm (Fig 5E). Consistent with these results, the association of DHX9 with rRNA genes in NIH 3T3 cells decreased upon depletion of TIP5 (Fig 5F), suggesting that DHX9 requires the association of TIP5 with mature pRNA to bind to rRNA genes. Finally, NIH 3T3 cells depleted of DHX9 showed a reduction in rRNA gene silencing. Upon DHX9

Published online: June 6, 2017
EMBO reports

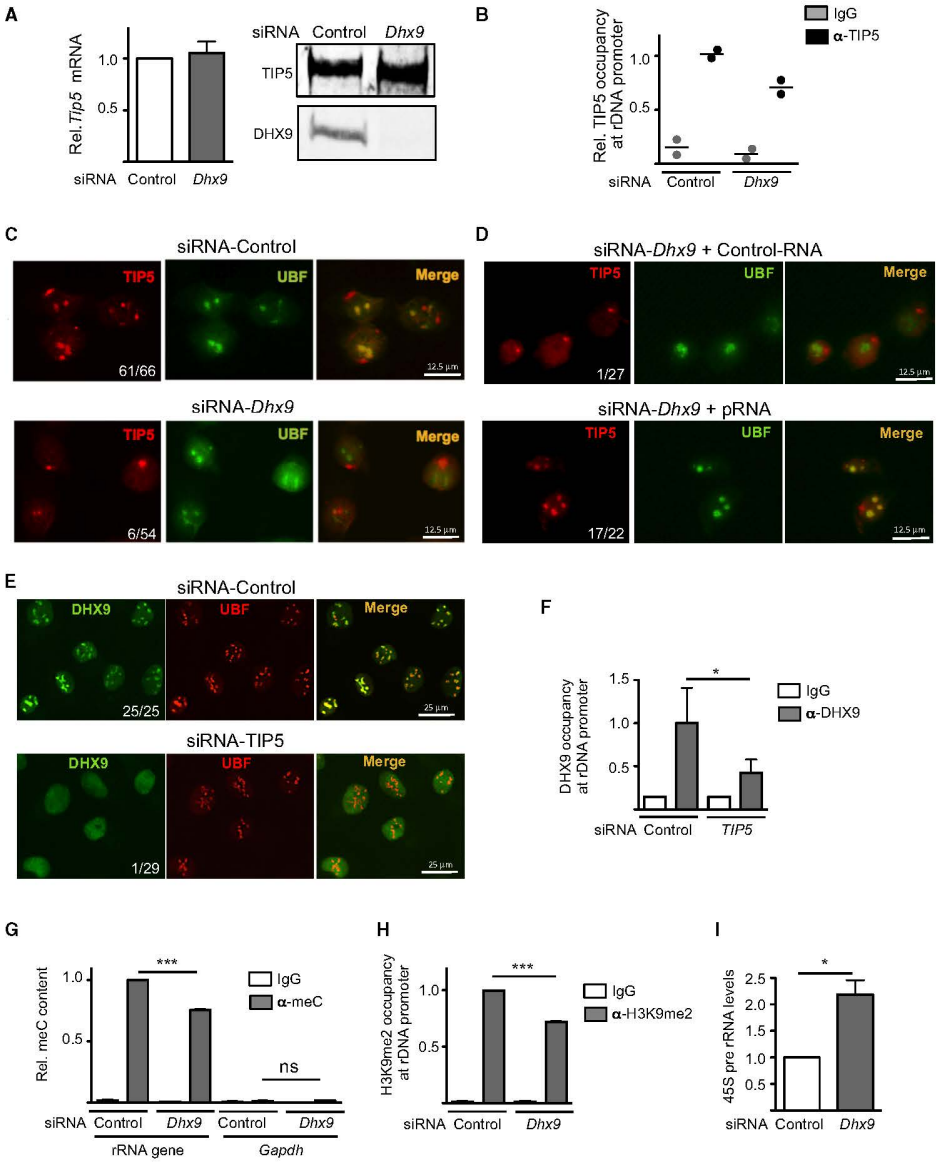


Figure 5.

Published online: June 6, 2017

Figure 5. DHX9 mediates recruitment of TIP5 and establishment of rRNA gene heterochromatin through processing of IGS-rRNA into pRNA.

A DHX9 knockdown does not affect TIP5 levels. mRNA and protein levels of TIP5 in NIH 3T3 cells depleted of DHX9 by siRNA. TIP5 mRNA values were normalized to *GAPDH* mRNA and to siRNA-control cells. Values (mean ± SD) are from three independent experiments.

B DHX9 is required for the association of TIP5 with rRNA genes. ChIP analysis of TIP5 occupancy at rRNA genes in NIH 3T3 cells depleted of DHX9 by siRNA. Data from two independent experiments are represented as bound over input and normalized to siRNA-Control cells. Scatter plot represents the values and the mean of two independent experiments.

C DHX9 is required for the localization of TIP5 in nucleoli. Immunofluorescence with anti-TIP5 and anti-UBF of NIH 3T3 cells treated with siRNA-Control or-DHX9. Numbers refer to cells showing TIP5 nucleolar localization relative to the number of analyzed cells.

D Retention of TIP5 in nucleoli depends on DHX9-mediated production of mature pRNA. Immunofluorescence with anti-TIP5 and anti-UBF of DHX9-depleted NIH 3T3 cells transfected with Control-RNA or pRNA.

E Nucleolar localization of DHX9 depends on TIP5. Immunofluorescence with anti-DHX9 and anti-UBF of U2OS cells treated with siRNA-Control or siRNA-*TIP5*. Numbers refer to cells showing DHX9 nucleolar localization relative to the number of analyzed cells.

F TIP5 is required for the association of DHX9 with rRNA genes. ChIP analysis of DHX9 occupancy at rRNA genes in U2OS cells depleted of TIP5 by siRNA. Values (mean ± SD) from three independent experiments are represented as bound over input and normalized to siRNA-Control cells.

G DHX9 is required for rRNA gene silencing. Methylated DNA immunoprecipitation (MeDIP) analysis of rRNA gene promoter using anti-5mC antibodies in NIH 3T3 cells upon DHX9 knockdown. Enrichments were calculated relative to input and normalized to rRNA genes in control cells. Low enrichment of GAPDH sequences (free of CpG methylation) ensures for the specificity of the measurement. Values (mean ± SD) are from three independent experiments.

H ChIP analysis of H3K9me2 at rRNA gene promoter in NIH 3T3 cells upon DHX9 knockdown. Enrichments were calculated relative to input and normalized to control cells. Values (mean ± SD) are from three independent experiments.

I Knockdown of DHX9 upregulates rRNA transcription. RT–qPCR of 45S pre-rRNA levels in NIH 3T3 cells upon DHX9 knockdown. Values (mean ± SD) from three independent experiments were normalized to *GAPDH* mRNA and to control cells.

Data information: Statistical significance (*P*-values) for the experiments was calculated using the paired two-tailed *t*-test (**P* < 0.05; ****P* < 0.001; ns, non-significant).

knockdown, rRNA genes displayed decreased levels of repressive epigenetic marks such as CpG methylation and H3K9me2 and increased 45S pre-rRNA transcription to levels similar to what reported upon TIP5 knockdown in NIH 3T3 cells [42] (Figs 5G–I and EV1G). Taken together, these results support a role of DHX9 in guiding TIP5 to rRNA genes and subsequent establishment of nucleolar heterochromatin in a manner dependent on the production of mature pRNA.

DHX9 localizes at nucleoli only upon ESC differentiation

Processing of IGS-rRNA into pRNA is impaired in ESCs and activated only upon ESC differentiation, which coincides with the timing of *de novo* heterochromatin formation at rRNA genes [13]. To determine whether the inhibition of IGS-rRNA processing in ESCs is dependent on alterations in DHX9 expression or cellular localization, we analyzed and compared ESCs before and after 4 days of differentiation achieved upon leukemia inhibitory factor (LIF) withdrawal. DHX9 is expressed in both ESCs and differentiated cells, the latter showing a slight increase at mRNA and protein levels (Fig EV5A and B). Remarkably, ChIP analysis revealed that the association of DHX9 with rRNA genes was higher in differentiated cells than in ESCs (Fig 6A). Therefore, although DHX9 and TIP5 associate in ESCs (Fig EV1D), in the absence of mature pRNA this interaction is not sufficient to recruit DHX9 or TIP5 to rRNA genes. Consistent with these results, analysis of GFP-DHX9 in ESCs revealed its localization in nuclei without any evident enrichment in nucleoli (Fig 6B–D). In contrast, upon differentiation GFP-DHX9 drastically changed its position within the nucleus and all cells showed DHX9 with an exclusive localization within nucleoli and a weak signal in the rest of nucleoplasm (Fig 6C). The accumulation of DHX9 in nucleoli and its association with rRNA genes only upon ESC differentiation correlates well with the activation of IGS-rRNA processing into pRNA, which takes place in nucleoli [19]. However, how DHX9 is excluded from nucleoli of ESCs still remains unknown.

Figure 6. DHX9-mediated production of mature pRNA is required for ESC differentiation.

A DHX9 associates with rRNA genes only upon ESC differentiation. ChIP analysis using DHX9 antibodies in ESCs and cells after 4 days of differentiation. Values (mean ± SD) are from three independent experiments and calculated relative to input and to values of rRNA sequences in ESCs.

B–D DHX9 is not enriched in nucleoli of ESCs. (B) Schema depicts the experimental strategy to measure DHX9 localization in ESCs and after 3 days of differentiation. (C) Immunofluorescence analysis with GFP and UBF antibodies of ESCs transfected with GFP-DHX9. (D) Life cell imaging of ESCs transfected with GFP-DHX9 (right). Nucleoli can be visualized by phase contrast.

E–G DHX9 is required for ESC differentiation. (E) Schema depicts the experimental strategy to analyze differentiation of ESCs depleted of DHX9 by siRNA. (F) RT–qPCR analysis of DHX9 and pluripotency factors in ESCs upon DHX9 knockdown. Expression levels were normalized to *Rps12* mRNA amounts and control ESCs. Average and standard deviation ($n = 3$) are shown for each point. (G) Representative images of alkaline phosphatase staining of ESCs and cells after 3 days of differentiation.

H–J DHX9-mediated production of mature pRNA is required for ESC differentiation. (H) Schema depicts the experimental strategy to analyze differentiation of ESCs depleted of DHX9 by siRNA in the presence of mature pRNA. (I, J) Quantification and representative images of ESCs after 3 days of differentiation. Values (mean ± SD) represent number of differentiated cells from three independent experiments relative to control cells (siRNA-control + Control-RNA).

K The association of TIP5 with mature pRNA is required for ESC differentiation. Quantification and representative images of ESCs treated with siRNA-DHX9 and transfected with pRNA mutants and IGS-rRNA after 3 days of differentiation. Values represent number of differentiated cells to control differentiated cells (siRNA-control + Control-RNA). Data represent the mean of two independent experiments.

Data information: Statistical significance (*P*-values) was calculated using the paired two-tailed *t*-test (**P* < 0.05; ***P* < 0.01; ns, non-significant)

Published online: June 6, 2017
EMBO reports

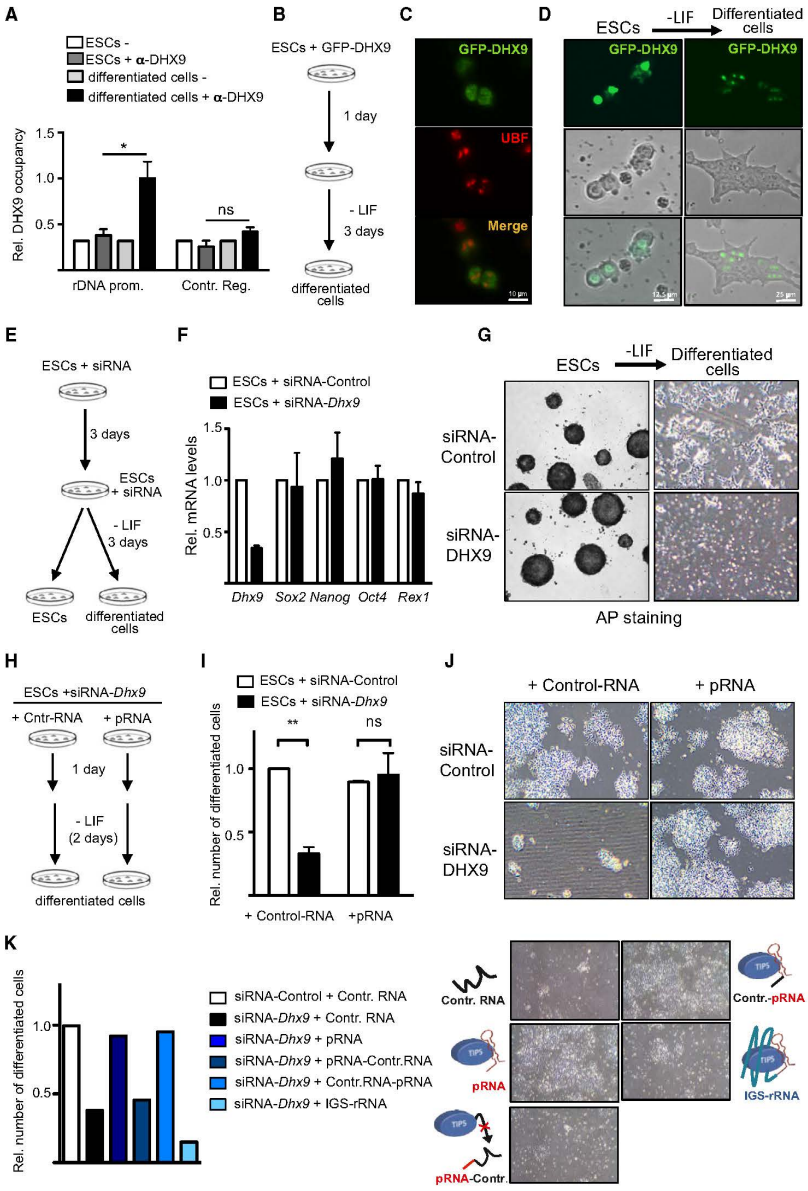


Figure 6.

Published online: June 6, 2017

DHX9-mediated processing of IGS-rRNA is required for ESC differentiation

Previous work has shown that during differentiation, the euchromatic and transcriptional permissive ESC genome remodels into a more repressive and heterochromatic structure [10,43]. Similarly, rRNA genes in ESCs are euchromatic and only upon ESC differentiation heterochromatin in the nucleolus is established [13]. Determinant for the lack of heterochromatin in nucleoli is the impairment of IGS-rRNA processing since addition of mature pRNA into ESCs was sufficient to induce heterochromatin formation at rRNA genes. Remarkably, the presence of ectopic pRNA in ESCs also initiated the establishment of highly condensed chromatin structures outside of the nucleolus, which was similar to what observed in differentiated cells. Moreover, such ESCs were primed for differentiation as indicated by the increased expression of genes implicated in developmental and differentiation processes. Therefore, we asked whether the production of mature pRNA mediated by DHX9 is necessary for ESC differentiation. Upon DHX9 knockdown, ESCs did not show alterations in important molecular features of the undifferentiated state such as cell proliferation, expression of the pluripotency genes, cell morphology, and alkaline phosphatase (AP) staining (Fig 6E–G). To determine whether DHX9 depletion affects ESC differentiation, we induced monolayer differentiation of ESCs treated with siRNA-control or siRNA-*Dhx9*. After 3 days, control cells displayed morphological structures typical of differentiated cells were negative for AP staining and had downregulation of pluripotency markers (Figs 6G and EV5C). In contrast, after induction of differentiation, cells depleted of DHX9 did not attach to the plate and underwent cell death (Fig 6G). To assess whether the requirement of DHX9 for ESC differentiation depends on its role in IGS-rRNA processing, we transfected mature pRNA into ESCs treated with siRNA-control or siRNA-*Dhx9* and subsequently induced differentiation 1 day later (Fig 6H). Both ESCs transfected with siRNA-control and with control-RNA or pRNA efficiently differentiated whereas ESCs depleted of DHX9 and transfected with control-RNA underwent cell death (Fig 6I and J). Remarkably, transfection of mature pRNA in ESCs knocked down for DHX9 was sufficient to entirely restore differentiation capability to same efficiencies as in control cells (Fig 6I and J). Consistent with the results described above, DHX9-depleted ESCs transfected with IGS-rRNA were unable to differentiate (Fig 6K), supporting the role of mature pRNA to drive cells into differentiation. Since the association of TIP5 with pRNA is key to establish heterochromatin at rRNA genes, we analyzed the ability of DHX9-depleted ESCs to differentiate using a pRNA mutant whose 3′-region necessary for TIP5 association was replaced with a control-RNA sequence (pRNA-Control) [13,15,16]. As shown in Fig 6K, DHX9-depleted ESCs transfected with pRNA-Control failed to differentiate. In contrast, replacement of the 5′-pRNA region (Control-pRNA), which is not required for the association of TIP5 and recruitment to rRNA genes [13], was sufficient to entirely restore ESC differentiation capacity, underscoring the importance of TIP5/pRNA-mediated heterochromatin formation at rRNA genes during ESC differentiation. Together, these results indicate that processing of IGS-rRNA into pRNA mediated by DHX9 is an event required for early differentiation and highlight the role of the chromatin state of rRNA genes in controlling this process.

Discussion

In this work, we have shown that the RNA helicase DHX9 is required to process IGS-rRNA into pRNA and this activity is required for ESC differentiation. Because of the role of pRNA to establish heterochromatin at rRNA genes and the fact that mature pRNA is produced only upon ESC differentiation, these results strongly support a role of the chromatin state of rRNA genes for the ability of ESCs to undergo differentiation.

The remodeling of the open and euchromatic genome structure of ESCs toward the formation of highly condensed heterochromatic structures marks the exit of ESCs from pluripotency and the entry into differentiation program [8]. This change in chromatin state is also characteristic of rRNA genes, a fraction of which only acquires epigenetic silent marks upon differentiation [13]. The formation of heterochromatin in the nucleolus during differentiation is also timely coincident with the clustering of highly condensed heterochromatin at nucleoli or at the nuclear periphery [11,12]. The link between the nucleolus and heterochromatin is also true for the inactive X-chromosome that contacts the nucleolus to duplicate its silent chromatin structures during mid-to-late S phase [44]. Similarly, lamina-associated domains (LADs) that are relatively gene poor and have a repressive chromatin signature were also found relocated next to the nucleolus [45,46]. Thus, by analogy with the nuclear periphery [47,48] also the clustering of heterochromatin at the nucleolus might play a role for the establishment of mammalian heterochromatin.

Our previous work has started to highlight the nucleolus as an important regulator that orchestrates the formation of heterochromatin. Indeed, inducing heterochromatin at rRNA genes in ESCs through addition of mature pRNA resulted in a drastic change of genome architecture with the appearance of highly condensed heterochromatic blocks outside the nucleolus [13]. These changes were also accompanied by a global increase in H3K9me2, maturation of heterochromatin at repetitive sequences—such as major and minor satellites—and their transcriptional repression as found in differentiated cells. Although the mechanisms through which the nucleolus acts in this global restructuring of genome architecture is yet unknown, what is clear is that it depends on the chromatin state of rRNA genes since the gain of heterochromatin at rRNA genes induces the rest of the genome to remodel into highly condensed structures.

The finding that DHX9 is required for processing of IGS-rRNA into pRNA, a reaction that is activated only upon ESC differentiation, allowed the determination that inhibition of this process impairs ESC differentiation. Indeed, the defects of differentiation observed upon DHX9 knockdown could be reverted solely by the addition of mature pRNA, underscoring the importance of DHX9 in the processing of IGS-rRNA into pRNA and subsequent establishment of nucleolar heterochromatin in early development. Remarkably, the requirement of DHX9 in early development is also supported by previous data indicating that DHX9 is essential for gastrulation and *Dhx9* $^{-/-}$ embryos do not progress further than E7.5 in development [49]. Moreover, the switch in the cellular localization of DHX9, which is lowly enriched in nucleoli of ESCs but exclusively present in nucleoli of differentiated cells, correlates well with the activation of IGS-rRNA processing that takes place in nucleoli only upon exit of pluripotency. However, how the cellular

Published online: June 6, 2017
EMBO reports

localization of DHX9 is regulated in early development is yet unclear and will be an aim of our future studies.

The contribution of DHX9-mediate production of pRNA in ESCs differentiation supports previous observations showing defects in ESC differentiation upon depletion of TIP5 [13]. Thus, impairment of rRNA gene heterochromatin achieved by alteration of pathways upstream TIP5 silencing activity without altering TIP5 expression levels brought to similar conclusion, that is the chromatin state of rRNA genes plays a role for the exit from pluripotency. This is further supported by the fact that pRNA mutants with impaired ability to associate with TIP5 failed to rescue the differentiation defects observed in ESCs upon depletion of DHX9. Recent studies have proposed that elevated rRNA biosynthesis sustains pluripotency in mouse and human ESCs since rRNA downregulation through chemical inhibition of Pol I activity or deletion of fibrillarin induces differentiation [21,22]. Although the set-up of our experiments does not allow us to determine whether the impairment of ESC differentiation is due to the maintenance of high rRNA transcription or to the lack of heterochromatin in nucleolus, we favor the latter case. Indeed, it is unlikely that cells with high rRNA level undergo cell death since ribosome biogenesis is well known to be positively correlated with cell viability and proliferation [1]. Thus, our data favor a model in which the nucleolus is not only the cellular compartment where ribosomes are produced but it is also a central component of nuclear architecture that coordinates the balance between euchromatin and heterochromatin according to developmental stages. Likewise, the function of rRNA genes might not only be limited to the synthesis of rRNA, a model which is in agreement with early studies showing that the fraction of silent rRNA genes present in each differentiated cell does not change its transcriptional state even under conditions of high metabolic activities [4]. The link between rRNA genes and the chromatin architecture of the rest of the genome is also supported by previous results in *Drosophila* showing that deletion of rRNA repeats reduced heterochromatin content elsewhere in the genome [50]. A similar observation was also found upon knockdown of TIP5 in NIH 3T3 cells, which induced not only a decrease of rRNA gene silencing but also the loss of perinucleolar heterochromatic blocks and the reduction of silent histone marks at pericentric heterochromatin [42].

Our data suggested that the splicing pathway might be implicated in the processing of IGS-rRNA into pRNA. The spliceosome was indeed the major hit obtained through cellular component and pathway analysis of both IGS-rRNA pulled-down proteins and DHX9-associated factors from ESCs, differentiated ESCs, NIH 3T3, and HEK293 cells. Remarkably, spliceosome core components (i.e., U2 Small Nuclear RNA Auxiliary Factor 2 U2AF2, Splicing factor 3B subunit 3 SF3B3, pre-mRNA Processing Factor 3 and 8, PRP3 and 8) have been found in the analysis of the proteome of nucleoli [51]. The possibility of a role of the spliceosome in the production of mature pRNA is also supported by the absence of ribonucleases among the identified IGS-rRNA interacting proteins. Xrn2 (5′–3′ exoribonuclease 2) was the unique ribonuclease we identified as DHX9-associated proteins in three out of four experiments. However, analyses of IGS-rRNA processing in cells depleted of Xrn2 by siRNA did not reveal any substantial defect as the ones found upon DHX9 knockdown (data not shown). Thus, the strong interaction of DHX9 with spliceosome components and the lack of association with ribonucleases with function in IGS-rRNA processing suggest a mechanism where DHX9 through binding to TIP5 and/or IGS-rRNA recruits the spliceosome complex for IGS-rRNA processing. The mechanisms of this reaction are currently under investigation.

Taken together, the results provided molecular insights into the biogenesis of lncRNAs that modulates different features of the same lncRNA according to developmental stage. Moreover, the data supported a model in which the state of rRNA gene chromatin is part of the regulatory network that controls the exit from pluripotency and the initiation of differentiation pathways.

Materials and Methods

Cell culture

NIH 3T3, U2OS, and HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco) and 1% penicillin/streptomycin (Gibco).

One hundred and twenty-nine mouse embryonic stem cells (E14 line) were cultured in DMEM-F12 and Neurobasal medium (1:1, Gibco), supplemented with 1× N2/B27 (Gibco), 1× penicillin/streptomycin/L-glutamine (Gibco), 50 μM β-mercaptoethanol (Gibco), recombinant leukemia inhibitory factor, LIF (ESGRO, 1,000 U/ml) and MEK and GSK3β inhibitors, 2i (Stemolecule CHIR99021 and PD0325901, 3 and 1 μM, respectively). ESCs were seeded at a density of 50,000 cells/cm^2 in culture dishes (Corning® CellBIND® surface) treated with 0.1% gelatine without feeder layer. Propagation of cells was carried out every 2 days using trypsin 0.5× for enzymatic cell dissociation. ESCs were differentiated by culturing for 48–72 h in complete medium: DMEM, 10% FCS, 1 mM sodium piruvate (Sigma), 1× NEAA (Gibco), 1× penicillin/streptomycin/L-glutamine, 100 μM β-mercaptoethanol.

The differentiation toward neural progenitor cells (NPC) was obtained according to previously established protocol [52]. In brief, differentiation employed a suspension-based embryoid bodies formation (Bacteriological Petri Dishes, Bio-one with vents, Greiner®) in neural differentiation media (DMEM, 10% FCS, 1× MEM NEAA, 1× penicillin/streptomycin/L-glutamine, 100 μM β-mercaptoethanol). During the 8-day differentiation procedure, media were exchanged every 2 days. In the last 4 days of differentiation, the media were supplemented with 2 μM retinoic acid (RA) to generate neuronal precursors (i.e., Pax-6-positive radial glial cells).

Transfections

Plasmids and siRNAs were transfected in HEK293T cells using calcium phosphate protocol. NIH 3T3 mouse fibroblasts, U2OS and ESCs were transfected with the indicated siRNAs (50 nM siRNA) using Lipofectamine® RNAiMAX (Life Technologies) in Opti-MEM® GlutaMAX™ (GIBCO) reduced-serum medium. Analysis of differentiated transfected ESCs was performed using consecutive transfections. To test ESC ability to differentiate upon DHX9 knockdown, 3 days after the first transfection, equal amounts of ESCs (e.g., siRNA-control and siRNA-*DHX9* treated cells) were again transfected and induced to differentiate in complete media (G-MEM, 10% FCS, sodium pyruvate 100 mM, 1× MEM NEAA, L-glutamine) by

Published online: June 6, 2017

withdrawal of LIF and 2i. Transfections of siRNA together with synthetic RNAs (1 mg/ml) were performed using Transit-X2® transfection reagent (Mirus), and the differentiation was induced 1 day after transfection by withdrawal of LIF and 2i.

Efficiencies of siRNA-mediated depletions and synthetic RNA levels were monitored by qRT–PCR 3–4 days post-transfection.

RNA extraction, reverse transcription, and quantitative PCR (RT–qPCR)

RNA was purified with TRIzol reagent (Life Technologies). Residual contaminating genomic DNA was removed with Ambion® TURBO™ DNase according to manufacturer's instructions. RNA was primed with random hexamers and reverse-transcribed to cDNA. Amplification of samples without reverse transcriptase assured absence of genomic or plasmid DNA (data not shown). The relative transcription levels were determined by normalizing to *Rps12* or *GAPDH* mRNA levels, as indicated. Measurements of IGS-rRNA processing were performed as previously described [53]. Reverse transcription was obtained using primers −20/−1 Rev (endogenous transcripts) or Vector RT primer hybridizing vector sequences (reporter assay). Processing efficiency was calculated by normalizing amplifications of the 5′-IGS-rRNA region (endogenous: primers −1,994/−1,975 For and −1,922/−1,905 Rev; reporter: primers −1,994/−1,975 For and Vector 1 Rev) to amplification of pRNA sequences (endogenous: primers −165/−145 For and −39/−20 Rev; reporter: primers −165/−145 For and Vector 2 Rev). qRT–PCR was performed with SensiMix SYBR Hi-ROX Mix (Bioline) on a Rotor-Gene Q (Qiagen). Statistical significance (*P*-values) of the difference in expression levels between genes was calculated using the two-sample paired *t*-test. Primer sequences are listed in Table EV1.

Chromatin immunoprecipitation (ChIP)

ChIP analysis was performed as previously described [54]. Briefly, formaldehyde 1% was added to cultured cells to cross-link proteins to DNA. Isolated nuclei were then lysed and sonicated using a Bioruptor ultrasonic cell disruptor to shear genomic DNA to an average fragment size of 200 bp. 20 μg of chromatin was diluted to a total volume of 500 μl with IP buffer (16.7 mM Tris–HCl, pH 8.1, 167 mM NaCl, 1.2 mM EDTA, 0.01% SDS, 1.1% Triton X-100) and pre-cleared with 10 μl packed Sepharose beads for 2 h at 4°C. Pre-cleared chromatin was incubated overnight with the indicated antibodies. The next day, Dynabeads protein-A (or -G, Millipore) were added and incubated for 4 h at 4°C. After washing, bound chromatin was eluted with the elution buffer (1% SDS, 100 mM $NaHCO_3$). Upon proteinase K digestion (50°C for 1 h) and reversion of cross-linking (65°C, overnight), DNA was purified with phenol/chloroform, ethanol precipitated and quantified by qPCR using the primers listed in Table EV1.

Methylated DNA immunoprecipitation (MeDIP)

200-bp length DNA fragments were denatured and incubated overnight with 5mC-antibody (Diagenode) in 200 μl IP buffer (10 mM Na-phosphate buffer, pH 7.0, 140 mM NaCl, 0.05% Triton X-100). The day after, 20 μl of protein-G Dynabeads (Millipore) was added and incubated for 2 h at 4°C. Beads were then washed three times at RT for 10 min with IP buffer and subsequently incubated with 250 μl proteinase K digestion mix (50 mM Tris–HCl, pH 8, 0.5% SDS, 10 mM EDTA, 2 μg proteinase K) overnight at 42°C. DNA was purified with phenol/chloroform, ethanol precipitated and quantified by qPCR.

AP staining

Cells were fixed in 4% paraformaldehyde for 10 min, washed with AP buffer (100 mM Tris-Cl pH 9.5, 100 mM NaCl, 50 mM MgCl2), and then incubated for 5–10 min in BCIP®/NBT liquid substrate system (Sigma). The staining was blocked with 10 mM Tris and 1 mM EDTA for 10 min.

GRNA chromatography

GRNA chromatography was performed as previously described [23]. Briefly, Glutathione Sepharose matrix was incubated for 1 h rotating at 4°C with GST-λ phage N antiterminator fusion protein and equimolar amounts of *in vitro* synthesized BoxB-RNAs (i.e., BoxB-Control-RNA and BoxB-IGS-rRNA) to obtain the RNA affinity matrix (GRNA resins). The GRNA resins were incubated overnight rotating at 4°C with nuclear extracts from NIH 3T3 cells or NPC in BB buffer (50 mM Tris–HCl pH 7.5, 100 mM KCl, 2 mM $MgSO_2$, 0.1% NP-40, 0.1 mg/ml yeast tRNA, 0.01 mg/ml heparin, 1× protease inhibitor cOmplete EDTA-free Roche), followed by extensive washing and elution of bound proteins using 20 μg/ml of RNase A (Thermo Scientific). Eluted proteins were precipitated in 20% trichloroacetic acid, washed extensively with pure acetone, and dried at 95°C for 5–10 min. The samples were solubilized and incubated overnight at 37°C in 30 μl of buffer (10 mM Tris, pH 8.2, 2 mM $CaCl_2$), 10 μl acetonitrile, 5 μl trypsin (100 ng/μl in 10 mM HCl), and 5 μl of RapiGest™ (1% in water). After centrifugation, the supernatants were dried and resuspended in 0.1% formic acid and analyzed by LC-MS/MS.

RNA immunoprecipitation

RNA immunoprecipitation was performed as previously described [55]. Briefly, 1% final concentration formaldehyde was added to the cell medium to cross-link proteins to RNA. Nuclei were then isolated in swelling buffer (5 mM HEPES pH 8.0, 85 mM KCl, 0.5% Nonidet P-40, 1× protease inhibitor cOmplete EDTA-free Roche) and lysed in nuclei lysis buffer (50 mM Tris–HCl pH 8.0, 10 mM EDTA pH 8.0, 1% w/v SDS, 1× protease inhibitor, 40 U/ml RNase inhibitor). The extracts were diluted tenfold with FA lysis buffer (1 mM EDTA pH 8.0, 50 mM HEPES pH 7.5, 140 mM NaCl, 0.1% w/v sodium deoxycholate, 1% v/v Triton X-100, 1× protease inhibitor, 40 U/ml RNase inhibitor) and sonicated for 5 min with 30 s on/off cycles in a Bioruptor sonicator (Diagenode). The extracts were pre-cleared with 20 μl packed Sepharose beads for 1 h rotating at 4°C. Pre-cleared extracts were then adjusted to 25 mM $MgCl_2$, and 5 mM $CaCl_2$ and incubated at 37°C for 30 min. with 120 μg/ml of DNase I (Fermentas). 5% of extracts was retained as input. Pre-cleared extracts were incubated overnight with the DHX9 antibody (Abcam ab26271) or GFP Trap® beads (Chromotek) to IP the endogenous or GFP-tagged overexpressed protein, respectively. The next day, in case of endogenous protein, Dynabeads protein-A (Millipore) were added

Published online: June 6, 2017
EMBO reports

and incubated for 4 h at 4°C. After stringent washing, bound RNA–protein complexes were eluted with RIP elution buffer (10 mM EDTA, 100 mM Tris–HCl pH 8.0, 1% SDS, 40 U/ml RNase inhibitor). Upon proteinase K digestion (42°C for 1 h) and reversion of cross-linking (65°C, 1 h), RNA was purified with an equal volume of acid-equilibrated phenol/chloroform and isopropanol precipitated. Residual contaminating genomic DNA was removed with Ambion® TURBO™ DNase according to manufacturer's instructions. After isopropanol precipitation, RNA was primed with random hexamers and reverse-transcribed to cDNA (as described above) and quantified by qPCR using the primers listed in Table EV1.

Immunoprecipitation

Nuclei were obtained by resuspending cells in hypotonic buffer (0.5% NP-40, 85 mM KCl, 5 mM HEPES, pH 7.4) and subsequent centrifugation at 6,000 *g* for 10 min at 4°C. Nuclei were then resuspended in nuclear extraction buffer (50 mM Tris–HCl pH 7.5, 0.15 M KCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 20% glycerol, 0.5 mM DTT, 0.5% NP-40 and 1× protease inhibitor cOmplete EDTA-free Roche) and sonicated two times for 30 s with a Bioruptor ultrasonic cell disruptor. After DNase I treatment for 1 h at 4°C, extracts were sonicated again for 30 s and centrifuged for 10 min at 4°C with 3,400 *g*. 0.5 mg of proteins from the resulting supernatant was immunoprecipitated overnight at 4°C using anti-HA-Agarose (Sigma) or M2 beads (Sigma). 0.05 mg (10%) of the extracts was later used to check for equal input material. Immunoprecipitates were washed three times at 4°C for 5 min with wash buffer (20 mM Tris–HCl, pH 7.8, 150 mM KCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 10% glycerol, 0.1% Tween and 1× protease inhibitor cOmplete EDTA-free Roche). Beads were collected by centrifuging for 5 min at 4°C with 500 rcf. Proteins were denatured with 1× Lämmli buffer, separated on SDS–polyacrylamide gel, and analyzed by immunoblotting.

For the analysis of DHX9 interactome, nuclei were obtained by resuspending FLAG-DHX9 expressing cells in nuclei preparation buffer (10 mM HEPES pH7.6; 1.5 mM $MgCl_2$; 10 mM KCl; 0.5 mM DTT and 1× protease inhibitor cOmplete EDTA-free Roche). After 10-min (HEK293T, ESC and ESC differentiated cells) or 30-min (NIH 3T3) incubation, cells were centrifuged at 3,800 *g* for 5 min, resuspended in nuclei preparation buffer, and subjected to 20–40 strokes in a dounce homogenizer. Nuclei were collected by centrifugation at 6,000 rcf, resuspended in nuclear extraction buffer, and immunoprecipitation was performed as described above. Elution of DHX9-interacting proteins was obtained by competition with 100 μg/ml of free FLAG peptide in wash buffer. Proteins of each elution were precipitated with trichloroacetic acid as described above for GRNA chromatography and analyzed by LC-MS/MS.

Immunofluorescence

TIP5 immunofluorescence was performed as described in ref. [13]. Briefly cells were seeded on glass coverslips 72–96 h before analysis and in the case of knockdown experiments transfected with siRNAs as described above. To perform the immunofluorescence, coverslips were washed with 1× PBS and cells were incubated with permeabilization buffer (20 mM Tris, pH 8, 5 mM $MgCl_2$, 0.5 mM EDTA, 25% glycerol, 0.05% Triton X-100) for 4 min at RT. After three washes with 1× PBS, cells were fixed with ice-cold pure methanol for 7 min at −20°C, washed again with 1× PBS, and incubated overnight with anti-TIP5 and anti-UBF. After washing three times with 1× PBS, cells were incubated with fluorescently labeled secondary antibodies for 1 h at RT, washed again, and stained with DAPI. Immunofluorescent images were digitally recorded.

DHX9 immunofluorescence was performed according to the protocol indicated by the DHX9 antibody manufacturer (Abgent). Briefly cells were seeded on glass coverslips 72–96 h before analysis. To perform the immunofluorescence, the coverslips were washed with 1× PBS and cells were fixed and permeabilized in 1× PBS containing 4% paraformaldehyde and 0.1% Triton X-100 for 20 min at RT followed by an incubation in blocking solution (1% BSA in PBS) for 30–60 min. After washing three times with 1× PBS, cells were incubated with anti-DHX9 for 1–3 h at RT, washed three times with 1× PBS, and incubated with fluorescently labeled secondary antibodies for 1 h at RT, washed again, and stained with DAPI. Immunofluorescent images were digitally recorded.

In vitro RNA synthesis

The following RNA and control sequences were cloned by PCR into pJET1.2 plasmids: Control-RNA (pJET1.2 backbone); pRNA (mrDNA from −232 to −1); Spacer-rRNA (mrDNA from −1,994 to −1,905); Enhancer-rRNA (mrDNA from −554 to −447); IGS-rRNA (mrDNA from −1,997 to +1). pRNA mutants (pRNA-Control and Control-pRNA were described in ref. [13]. BoxB-Control (BoxB sequence fused to pJET1.2 backbone from 345 to 2,134) and BoxB-IGS-rRNA (BoxB sequence fused to mrDNA from −1,994 to −1) were cloned by PCR into pTOPO 2.1 plasmid. Synthetic RNAs were synthesized using T7 polymerase and as substrate XbaI (pJET1.2), BamHI (pTOPO 2.1), or AvaII (pJET1.2 BoxB-Control) linearized vectors containing the indicated sequences. IGS-rRNA used for ESCs transfection was synthesized using the HiScribe™ T7 ARCA Kit (Neb E2065S) according manufacturer's instructions, in order to increase its stability upon transfection. After treatment with DNase I, transcripts were double purified using TRIzol reagent (Invitrogen) according to the manufacture's protocol, quantified and analyzed by agarose gel electrophoresis.

EMSA competition assay

Radiolabeled RNA_{MCS} was synthesized by T7 RNA polymerase using EcoRI linearized pBluescript-KS(+) plasmid as template. After treatment with DNase I, transcripts were purified and 50,000 cpm of MCS-RNA was incubated for 15 min on ice with 75 ng recombinant TIP5 or 50 ng DHX9 in EMSA buffer (20 mM Tris–HCl pH 8.0, 5 mM $MgCl_2$, 100 mM KCl, and 0.2 mM EDTA). The amounts of TIP5 and DHX9 moieties to be used in EMSA competition assay were determined by pilot titration experiments as the minimal amount of proteins necessary to obtain a complete shift of radiolabeled RNA_{MCS} in the absence of competitor RNA (shown in the second lane of each EMSA competition experiment). Cold competitor RNA was added, and incubation was continued for 30 min. RNA–protein complexes were analyzed by electrophoresis on 6% (w/v) native polyacrylamide gels and visualized by autoradiography. To produce recombinant DHX9, HEK 293T cells overexpressing FLAG-DHX9-His were resuspended in lysis buffer (10 mM Tris–HCl pH 7.5, 500 mM NaCl, 5 mM $MgCl_2$, 15 mM imidazole, 10%

Published online: June 6, 2017
Sergio Leone et al DHX9 regulates lncRNA processing EMBO reports

glycerol, 2 mM β-mercaptoethanol, 1× protease inhibitor cOmplete EDTA-free Roche), incubated for 10 min at 4°C, and sonicated two times for 30 s with a Bioruptor ultrasonic cell disruptor. Extracts were then treated with DNase I (10 U/ml, Thermo Scientific) and RNase A (40 μg/ml, Thermo Scientific) for 1 h at 4°C and centrifuged at 20,000 *g* for 30 min. Supernatant was incubated with ProBond™ resin (Invitrogen) at 4°C for 3 h to capture His-tagged proteins. After washing using lysis buffer, DHX9-bound bead were resuspended in lysis buffer and treated again with DNase I and RNase A as described above. After extensive washing using lysis buffer containing 500 mM NaCl, DHX9 was eluted using lysis buffer containing 300 mM NaCl and 300 mM imidazole. To produce recombinant TIP5, the N-terminal region of TIP5 fused to His tag ($TIP5_{1-600}$-His), which contains the RNA-binding domain TAM, was expressed in *Escherichia coli* BL21. Cell pellets were resuspended in lysis buffer (20 mM Tris–HCl, pH 8, 300 mM KCl, 5 mM imidazole, 1 mM 2-mercaptoethanol, and 1× protease inhibitor cOmplete EDTA-free Roche), treated with 50 U DNaseI for 30 min, and centrifuged for 30 min at 20,000 g. Supernatants were incubated with ProBond™ resin for 4 h, washed with wash buffer (50 mM Tris–HCl, pH 8, 500 mM KCl, 20 mM imidazole, 1 mM 2-mercaptoethanol and protease inhibitors), treated with RNase A, followed by further washing steps. $TIP5_{1-600}$ was eluted using wash buffer containing 250 mM imidazole. The quality and identity of purified proteins were assessed by Coomassie Brilliant Blue staining and Western blot analysis.

Antibodies

The following antibodies were used: anti-DHX9 (AW5241) from Abgene; anti-5mc (MAb-081-100) and anti-TIP5 (CS-090-100) from Diagenode; anti-UBF (sc-13125) and anti-DHX9 (sc-66997) from Santa Cruz; anti-GFP (11814460001) from Roche; anti-H3K9me2 (17-648) from Millipore; anti-HA (MMS-101P) from Covance.

Expanded View for this article is available online.

Acknowledgements
This work was supported by the Swiss National Science Foundation (31003A_173056 and 310003A-152854), Krebsliga Zurich, Julius Müller Stiftung, Olga Mayenfisch Stifung, Stiftung für wissenschaftliche Forschung an der Universität Zürich, and Forschungskredit of the University of Zurich. We thank Kevin Czaplinski for plasmids and protocols for CRNA chromatography, Li Xing for DHX9 expressing plasmids, and Peter Hunziker and the Functional Genomic Center Zurich for the assistance in proteomic analysis. We also thank Ulrike Kutay, Michael O. Hottiger, Christian Lehner, and Davide Gabellini for helpful discussions during the progression of this work.

Author contributions
SL performed screening for IGS-rRNA interacting proteins, IGS-rRNA processing analysis, biochemical assays, immunofluorescence experiments, ESC differentiation experiments, contributed to experimental design and data analysis, and wrote the manuscript; DB performed EMSA assays; CFS performed FLAG-DHX9 IP and contributed to the analysis of the interactome; DD performed the co-IP for endogenous TIP5 and DHX9; RS conceived and supervised the project and wrote the manuscript.

Conflict of interest
The authors declare that they have no conflict of interest.

References

1. Moss T, Stefanovsky VY (2002) At the center of eukaryotic life. *Cell* 109: 545 548
2. Santoro R, Grummt I (2001) Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription. *Mol Cell* 8: 719 725
3. Santoro R, Li J, Grummt I (2002) The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nat Genet* 32: 393 396
4. Conconi A, Widmer RM, Koller T, Sogo JM (1989) Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. *Cell* 57: 753 761
5. Pombo A, Dillon N (2015) Three-dimensional genome architecture: players and mechanisms. *Nat Rev Mol Cell Biol* 16: 245 257
6. Fussner E, Ahmed K, Dehghani H, Strauss M, Bazett-Jones DP (2010) Changes in chromatin fiber density as a marker for pluripotency. *Cold Spring Harb Symp Quant Biol* 75: 245 249
7. Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* 10: 105 116
8. Gaspar-Maia A, Alajem A, Meshorer E, Ramalho-Santos M (2011) Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol* 12: 36 47
9. Gorkin DU, Leung D, Ren B (2014) The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell* 14: 762 775
10. Meshorer E, Misteli T (2006) Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* 7: 540 546
11. Bartova E, Galiova G, Krejci J, Harnicarova A, Strasak L, Kozubek S (2008) Epigenome and chromatin structure in human embryonic stem cells undergoing differentiation. *Dev Dyn* 237: 3690 3702
12. Wiblin AE, Cui W, Clark AJ, Bickmore WA (2005) Distinctive nuclear organisation of centromeres and regions involved in pluripotency in human embryonic stem cells. *J Cell Sci* 118: 3861 3868
13. Savic N, Bar D, Leone S, Frommel SC, Weber FA, Vollenweider E, Ferrari E, Ziegler U, Kaech A, Shakhova O *et al* (2014) lncRNA maturation to initiate heterochromatin formation in the nucleolus is required for exit from pluripotency in ESCs. *Cell Stem Cell* 15: 720 734
14. Schlesinger S, Selig S, Bergman Y, Cedar H (2009) Allelic inactivation of rDNA loci. *Genes Dev* 23: 2437 2447
15. Guetg C, Scheifele F, Rosenthal F, Hottiger MO, Santoro R (2012) Inheritance of silent rDNA chromatin is mediated by PARP1 via noncoding RNA. *Mol Cell* 45: 790 800
16. Mayer C, Schmitz KM, Li J, Grummt I, Santoro R (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol Cell* 22: 351 361
17. Zhou Y, Santoro R, Grummt I (2002) The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. *EMBO J* 21: 4632 4640
18. Mayer C, Neubert M, Grummt I (2008) The structure of NoRC-associated RNA is crucial for targeting the chromatin remodelling complex NoRC to the nucleolus. *EMBO Rep* 9: 774 780

Published online: June 6, 2017
EMBO reports

19. Santoro R, Schmitz KM, Sandoval J, Grummt I (2010) Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans. *EMBO Rep* 11: 52 58
20. Grummt I, Kuhn A, Bartsch I, Rosenbauer H (1986) A transcription terminator located upstream of the mouse rDNA initiation site affects rRNA synthesis. *Cell* 47: 901 911
21. Watanabe-Susaki K, Takada H, Enomoto K, Miwata K, Ishimine H, Intoh A, Ohtaka M, Nakanishi M, Sugino H, Asashima M *et al* (2014) Biosynthesis of ribosomal RNA in nucleoli regulates pluripotency and differentiation ability of pluripotent stem cells. *Stem Cells* 32: 3099 3111
22. Woolnough JL, Atwood BL, Liu Z, Zhao R, Giles KE (2016) The regulation of rRNA gene transcription during directed differentiation of human embryonic stem cells. *PLoS One* 11: e0157276
23. Czaplinski K, Kocher T, Schelder M, Segref A, Wilm M, Mattaj IW (2005) Identification of 40LoVe, a *Xenopus* hnRNP D family protein involved in localizing a TGF-beta-related mRNA during oogenesis. *Dev Cell* 8: 505 515
24. Andersen JS, Lyon CE, Fox AH, Leung AK, Lam YW, Steen H, Mann M, Lamond AI (2002) Directed proteomic analysis of the human nucleolus. *Curr Biol* 12: 1 11
25. Fuchsova B, Hozak P (2002) The localization of nuclear DNA helicase II in different nuclear compartments is linked to transcription. *Exp Cell Res* 279: 260 270
26. Zhang S, Kohler C, Hemmerich P, Grosse F (2004) Nuclear DNA helicase II (RNA helicase A) binds to an F-actin containing shell that surrounds the nucleolus. *Exp Cell Res* 293: 248 258
27. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP *et al* (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 43: D447 D452
28. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 44: D457 D462
29. Jarmoskaite I, Russell R (2014) RNA helicase proteins as chaperones and remodelers. *Annu Rev Biochem* 83: 697 725
30. Zhang S, Grosse F (1994) Nuclear DNA helicase II unwinds both DNA and RNA. *Biochemistry* 33: 3906 3912
31. Zhang S, Grosse F (1997) Domain structure of human nuclear DNA helicase II (RNA helicase A). *J Biol Chem* 272: 11487 11494
32. Lee CC, Hurwitz J (1992) A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3′ to 5′ direction. *J Biol Chem* 267: 4398 4407
33. Kuroda MI, Kernan MJ, Kreber R, Ganetzky B, Baker BS (1991) The maleless protein associates with the X chromosome to regulate dosage compensation in *Drosophila*. *Cell* 66: 935 947
34. Myohanen S, Baylin SB (2001) Sequence-specific DNA binding activity of RNA helicase A to the p16INK4a promoter. *J Biol Chem* 276: 1634 1642
35. Anderson SF, Schlegel BP, Nakajima T, Wolpin ES, Parvin JD (1998) BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. *Nat Genet* 19: 254 256
36. Nakajima T, Uchida C, Anderson SF, Lee CG, Hurwitz J, Parvin JD, Montminy M (1997) RNA helicase A mediates association of CBP with RNA polymerase II. *Cell* 90: 1107 1112
37. Zhong X, Safa AR (2004) RNA helicase A in the MEF1 transcription factor complex up-regulates the MDR1 gene in multidrug-resistant cancer cells. *J Biol Chem* 279: 17134 17141
38. Aratani S, Fujii R, Oishi T, Fujita H, Amano T, Ohshima T, Hagiwara M, Fukamizu A, Nakajima T (2001) Dual roles of RNA helicase A in CREB-dependent transcription. *Mol Cell Biol* 21: 4460 4469
39. Jain A, Bacolla A, Del Mundo IM, Zhao J, Wang G, Vasquez KM (2013) DHX9 helicase is involved in preventing genomic instability induced by alternatively structured DNA in human cells. *Nucleic Acids Res* 41: 10345 10357
40. Chakraborty P, Grosse F (2011) Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair (Amst)* 10: 654 665
41. Xing L, Liang C, Kleiman L (2011) Coordinate roles of Gag and RNA helicase A in promoting the annealing of formula to HIV-1 RNA. *J Virol* 85: 1847 1860
42. Guetg C, Lienemann P, Sirri V, Grummt I, Hernandez-Verdun D, Hottiger MO, Fussenegger M, Santoro R (2010) The NoRC complex mediates the heterochromatin formation and stability of silent rRNA genes and centromeric repeats. *EMBO J* 29: 2135 2146
43. Bhattacharya D, Talwar S, Mazumder A, Shivashankar GV (2009) Spatio-temporal plasticity in chromatin organization in mouse cell differentiation and during *Drosophila* embryogenesis. *Biophys J* 96: 3832 3839
44. Zhang LF, Huynh KD, Lee JT (2007) Perinucleolar targeting of the inactive X during S phase: evidence for a role in the maintenance of silencing. *Cell* 129: 693 706
45. Kind J, Pagie L, Ortabozkoyun H, Boyle S, de Vries SS, Janssen H, Amendola M, Nolen LD, Bickmore WA, van Steensel B (2013) Single-cell dynamics of genome-nuclear lamina interactions. *Cell* 153: 178 192
46. Kind J, van Steensel B (2010) Genome-nuclear lamina interactions and gene regulation. *Curr Opin Cell Biol* 22: 320 325
47. Pinheiro I, Margueron R, Shukeir N, Eisold M, Fritzsch C, Richter FM, Mittler G, Genoud C, Goyama S, Kurokawa M *et al* (2012) Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. *Cell* 150: 948 960
48. Towbin BD, Gonzalez-Aguilera C, Sack R, Gaidatzis D, Kalck V, Meister P, Askjaer P, Gasser SM (2012) Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* 150: 934 947
49. Lee CG, da Costa Soares V, Newberger C, Manova K, Lacy E, Hurwitz J (1998) RNA helicase A is essential for normal gastrulation. *Proc Natl Acad Sci USA* 95: 13709 13713
50. Paredes S, Maggert KA (2009) Ribosomal DNA contributes to global chromatin regulation. *Proc Natl Acad Sci USA* 106: 17829 17834
51. Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M (2005) Nucleolar proteome dynamics. *Nature* 433: 77 83
52. Bibel M, Richter J, Schrenk K, Tucker KL, Staiger V, Korte M, Goetz M, Barde YA (2004) Differentiation of mouse embryonic stem cells into a defined neuronal lineage. *Nat Neurosci* 7: 1003 1009
53. Savić N, Bär D, Leone S, Frommel SC, Weber FA, Vollenweider E, Ferrari E, Ziegler U, Kaech A, Shakhova O *et al* (2014) lncRNA maturation to initiate heterochromatin formation in the nucleolus is required for exit from pluripotency in ESCs. *Cell Stem Cell* 720 734
54. Santoro R (2014) Analysis of chromatin composition of repetitive sequences: the ChIP-Chop assay. *Methods Mol Biol* 1094: 319 328
55. Selth LA, Close P, Svejstrup JQ (2011) Studying RNA-protein interactions *in vivo* by RNA immunoprecipitation. *Methods Mol Biol* 791: 253 264

Expanded View Figures

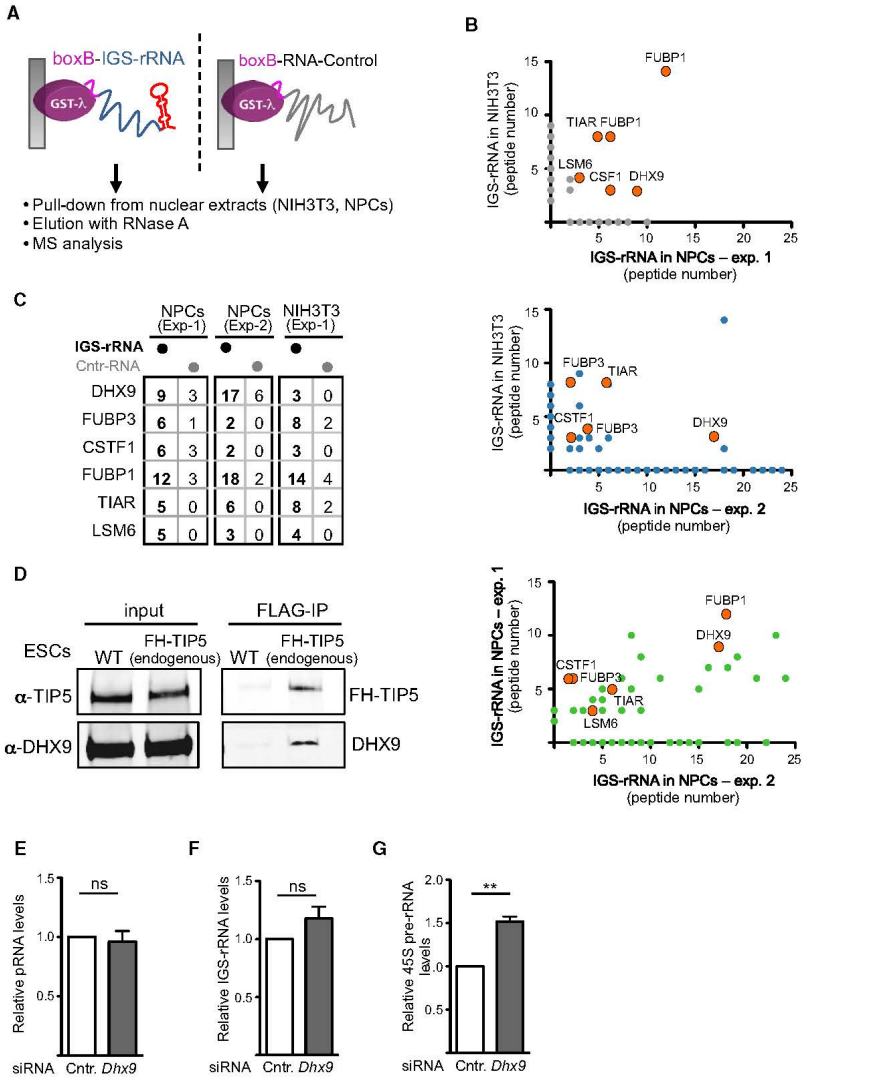


Figure EV1.

◀ **Figure EV1. DHX9 associates with IGS-rRNA and TIP5.**

A Schema represents the GRNA chromatography employed for the identification of IGS-rRNA-binding proteins.

B Mass spectrometry analysis of proteins pulled down with IGS-rRNA and found enriched in all three independent experiments, which used nuclear extracts from NIH 3T3 cells or NPCs. Values refer to peptide numbers obtained in each experiment (gray, blue and green circles). Orange circles highlight the six IGS-rRNA-binding proteins found to associate with IGS-rRNA in all experiments.

C Comparison of peptides pulled down with IGS-rRNA in three independent experiments using nuclear extracts from NIH 3T3 cells and NPCs. Values refer to peptide numbers obtained in each experiment.

D Co-IP showing the interaction of endogenous TIP5 and DHX9 in embryonic stem cells (ESCs). Flag-HA (FH)-TIP5 ESCs were obtained by co-transfection of the CRISPR/Cas9 vector expressing an sgRNA targeting the TIP5 genomic locus 3 nt upstream of the ATG start codon with a homology directed repair (HDR) template. Immunoblots of Flag-IP and input show the association of endogenous FH-TIP5 with DHX9.

E–G Knockdown of DHX9 upregulates 45S pre-rRNA levels but does not affect the total amount of IGS-rRNA and pRNA. qRT–PCR from NIH 3T3 cells transfected with siRNA-control and siRNA-DHX9. Values were obtained by amplification of cDNA generated in reverse transcription reactions using random hexamers. Values (mean ± SD) were from four independent experiments and calculated relative to *GAPDH* mRNA amounts and to control cells (siRNA-control). Statistical significance (*P*-values) for the experiments was calculated using the paired two-tailed *t*-test (***P* < 0.01; ns, non-significant).

Source data are available online for this figure.

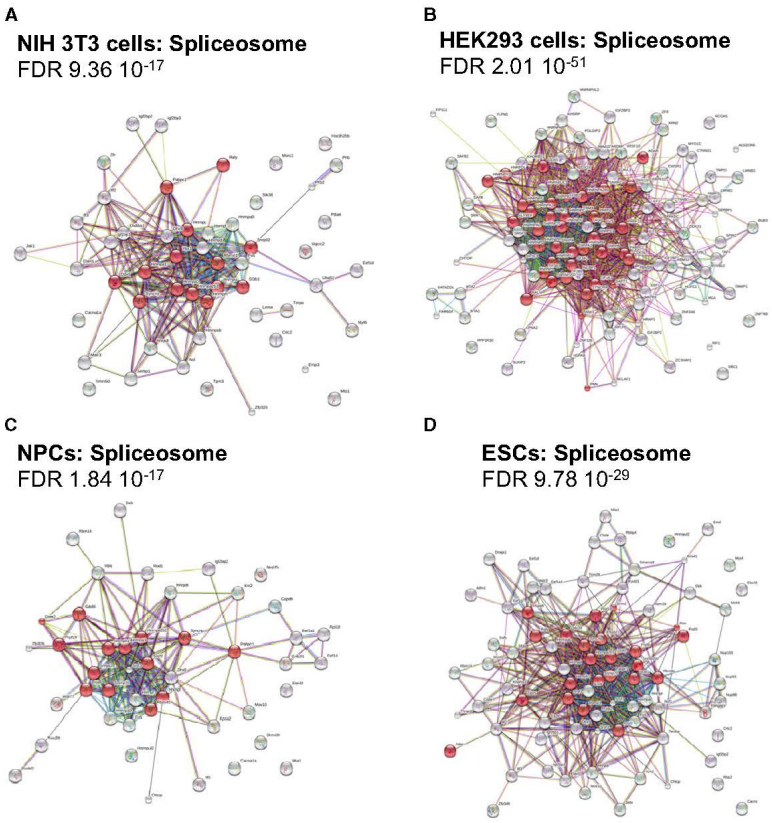


Figure EV2.

Figure EV2. DHX9 associates with components of RNA processing and spliceosome.

A–D STRING analysis depicting functional protein association networks of DHX9 interacting proteins found in NIH 3T3 cells (A), HEK293T (B), ESCs after 3-day differentiation (C), and ESCs (D). DHX9-interacting proteins were identified by transfection of FLAG-DHX9 plasmid, followed by anti-FLAG immuno-precipitation, FLAG peptide elution, and mass spectrometry analysis. Identified proteins are listed in Dataset EV1. Biological pathways and KEGG analysis are in Dataset EV2.

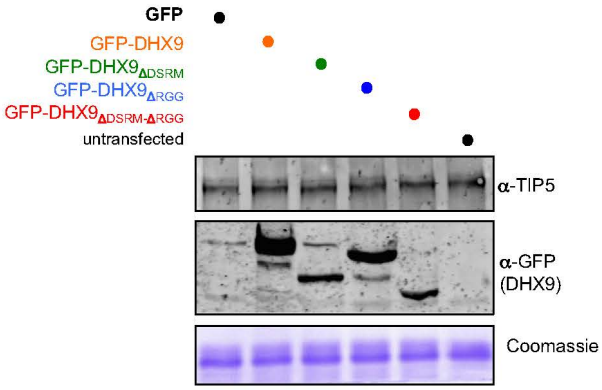


Figure EV3. Ectopic expression of DHX9 wild type and mutants does not affect TIP5 endogenous levels.

Western blot of lysates of HEK293 cells transfected with the indicated DHX9 expression vectors. Coomassie staining serves as control for equal loading between samples.

Source data are available online for this figure.

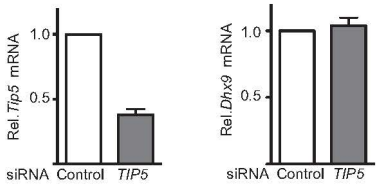


Figure EV4. TIP5 knockdown does not affect DHX9 levels.

mRNA of TIP5 and DHX9 in NIH 3T3 cells depleted of TIP5 by siRNA. TIP5 mRNA values were normalized to *GAPDH* mRNA and to siRNA-control cells. Average and standard deviation ($n = 3$) are shown for each point.

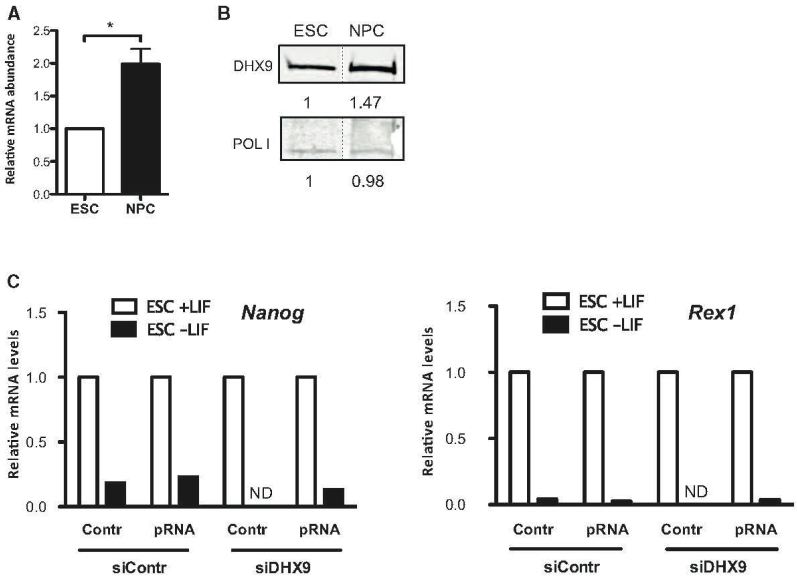


Figure EV5. Expression levels of DHX9 in ESCs and NPCs.

A, B DHX9 transcript (A) and protein (B) levels in ESCs and NPCs. RT–qPCR values (mean ± SD) from three independent experiments were normalized to *Rps12* mRNA levels and to ESCs. Statistical significance (*P*-values) for the experiments was calculated using the paired two-tailed *t*-test (*$P < 0.05$).

C qRT–PCR showing expression levels of pluripotency markers *Nanog* and *Rex1* in ESCs + LIF and differentiated cells obtained upon withdrawal of LIF (ESCs−LIF). Values were normalized to *rsp12* mRNA and the corresponding ESCs. Because of the few amounts of differentiated cells obtained upon DHX9 knockdown, mRNA levels in these samples could not be determined (ND). Data represent the mean of two independent experiments.

Source data are available online for this figure.

ARTICLES

nature cell biology

Pramel7 mediates ground-state pluripotency through proteasomal–epigenetic combined pathways

Urs Graf[1,2,3], Elisa A. Casanova[1], Sarah Wyck[3,4,5], Damian Dalcher[3,4], Marco Gatti[6], Eva Vollenweider[3,4], Michal J. Okoniewski[7], Fabienne A. Weber[2,3], Sameera S. Patel[2,3], Marc W. Schmid[8], Jiwen Li[9], Jafar Sharif[10], Guido A. Wanner[1], Haruhiko Koseki[10], Jiemin Wong[9], Pawel Pelczar[11], Lorenza Penengo[6], Raffaella Santoro[4,12,13] and Paolo Cinelli[1,2,12,13]

Naive pluripotency is established in preimplantation epiblast. Embryonic stem cells (ESCs) represent the immortalization of naive pluripotency. 2i culture has optimized this state, leading to a gene signature and DNA hypomethylation closely comparable to preimplantation epiblast, the developmental ground state. Here we show that Pramel7 (PRAME-like 7), a protein highly expressed in the inner cell mass (ICM) but expressed at low levels in ESCs, targets for proteasomal degradation UHRF1, a key factor for DNA methylation maintenance. Increasing Pramel7 expression in serum-cultured ESCs promotes a preimplantation epiblast-like gene signature, reduces UHRF1 levels and causes global DNA hypomethylation. Pramel7 is required for blastocyst formation and its forced expression locks ESCs in pluripotency. Pramel7/UHRF1 expression is mutually exclusive in ICMs whereas Pramel7-knockout embryos express high levels of UHRF1. Our data reveal an as-yet-unappreciated dynamic nature of DNA methylation through proteasome pathways and offer insights that might help to improve ESC culture to reproduce *in vitro* the *in vivo* ground-state pluripotency.

Embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of blastocysts and can be expanded indefinitely *in vitro*[1,2]. However, only ESCs have the capacity for unlimited self-renewal while ICM cells only briefly proliferate before acquiring a more restricted developmental potential[3,4]. Consequently, during the conversion of ICM cells to ESCs, there is an evident arrest of a normal developmental program, which *in vitro* is subverted in favour of a potential for unrestricted self-renewal while retaining pluripotency[5].

Depending on culture conditions, pluripotent ESCs can acquire molecular features that are distinct from those characterizing the developmental ground state of epiblast cells. Cultivation of ESCs in the presence of either serum/leukaemia inhibitory factor (LIF) or MEK/ERK and GSK3β inhibitors (2i, PD0325901 and CHIR99021) is conducive for maintenance of naive pluripotency. Most ESCs in serum/LIF (ESCs+serum) exhibit an altered transcriptional and epigenetic profile relative to preimplantation epiblast cells and are considered to be functionally naive but not ground state[6]. When compared to ICM, ESCs+serum have a globally more repressive epigenetic status as evident by a higher DNA methylation content and elevated expression of epigenetic regulators linked to transcriptional repression[5,7,8]. Conversely, a large proportion of epigenetic modifiers known to confer an active epigenetic state show higher expression in ICM.

Gene expression and DNA methylation analyses have shown that ESCs grown in 2i (ESCs+2i) are in an optimized state of naive pluripotency that closely resembles the developmental ground state *in vivo*[9–13]. Compared with ESCs+serum, ESCs+2i exhibit a permissive epigenetic landscape, including a hypomethylated genome that also characterizes ICM[7,8,10–12]. Thus, until now, the 2i-culture system represents the best available approach to model the developmental state of preimplantation epiblast cells in ESCs.

[1]Department of Trauma Surgery, Center for Clinical Research, University Hospital Zurich, University of Zurich, Sternwartstrasse 14, CH-8091 Zurich, Switzerland. [2]Institute of Laboratory Animal Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. [3]Life Science Zurich Graduate School, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. [4]Department of Molecular Mechanisms of Disease, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. [5]Clinic of Reproductive Medicine, University of Zurich, Winterthurerstrasse 260, CH-8057 Zurich, Switzerland. [6]Institute of Molecular Cancer Research, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. [7]Scientific IT Services, ETH Zurich, Weinbergstrasse 11, CH-8092 Zurich, Switzerland. [8]Service and Support for Science IT, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. [9]Shanghai Key Laboratory of Regulatory Biology, Institute of Biomedical Sciences and School of Life Sciences, East China Normal University, Shanghai 200241, China. [10]Developmental Genetics Laboratory, RIKEN Center for Integrative Medical Sciences, 1-7-22 Suehiuro-cho, Tsurumi-ku, Yokohama City, Kanagawa 230-0045, Japan. [11]Center for Transgenic Models, University of Basel, Mattenstrasse 22, CH-4002 Basel, Switzerland. [12]Center for Applied Biotechnology and Molecular Medicine, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland.
[13]Correspondence should be addressed to R.S. or P.C. (e-mail: raffaella.santoro@dmmd.uzh.ch or paolo.cinelli@usz.ch)

Received 9 November 2016; accepted 11 May 2017; published online 12 June 2017; corrected after print 21 June 2017; DOI: 10.1038/ncb3554

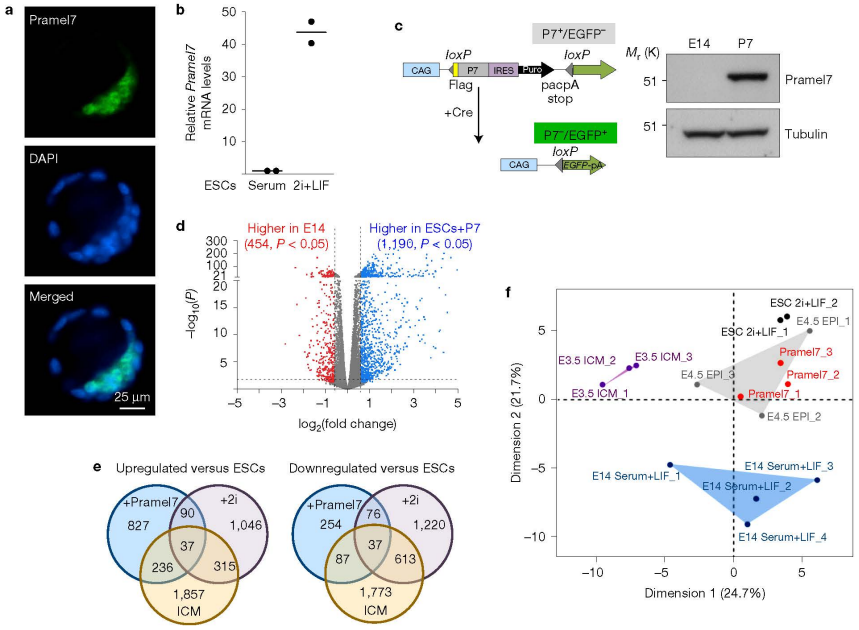


Figure 1 Expression of Prame17 in ESCs induces a gene signature similar to ICM. (**a**) Prame17 is expressed in ICM. Immunofluorescence analysis of a blastocyst showing expression of Prame17 in ICM but not in trophoblast cells. (**b**) *Prame17* gene expression in ESCs+2i compared with ESCs+serum. *Prame17* mRNA levels were measured by qRT–PCR and normalized to actin mRNA and to ESCs+serum. Average values of two independent experiments; lines represent means. (**c**) Scheme displaying the construct used to establish P7 ESCs, which expresses FLAG-Prame17 under the control of the CAG promoter and contains a Prame17-IRES-Puro cassette flanked by *loxP* sites followed by *EGFP* sequences (P7 ESCs). Western blot indicates expression levels of Prame17 in E14 ESCs and P7 ESCs (P7) detected with Prame17 antibodies. Tubulin is shown as a loading control. (**d**) Volcano plot showing fold change ($\log_2$ values) in transcript level of all coding genes in ESCs+Prame17 (P7) versus control ESCs (E14). Gene expression values of three replicates were averaged and selected for $P < 0.05$. A 1.5-fold change is indicated by the dotted line. (**e**) Venn diagrams showing transcript level changes detected in ESCs grown in serum compared with ESCs+Prame17, ICM (ref. 5) and ESCs grown in 2i (ref. 12). (**f**) Correlation of P7 ESC gene expression to the early embryo stages. PCA analysis of P7 ESCs, embryonic stages E3.5 and E4.5, ESCs cultured in serum/LIF or 2i from the data set published in ref. 9. The image shown in **a** is representative of three independent stainings. Unprocessed original scans of immunoblots are shown in Supplementary Fig. 6.

We and others have recently found that Prame17 (PRAME-like 7), a member of the Preferentially Expressed Antigen in Melanoma (PRAME) family, is expressed at high levels at the morula stage and in ICM, at lower levels in ESCs, but is completely absent in post-implantation embryos and in somatic tissues[14,15]. Correct dosage of Prame17 levels is critical in early development. In ESCs, *Prame17* knockdown induces differentiation, whereas its overexpression is sufficient to suppress differentiation in the absence of exogenous LIF, impairs teratoma formation and leads to embryonic lethality[15,16]. Moreover, expression of Prame17 in combination with the GSK3β-inhibitor CHIR99021 is sufficient to maintain ESCs undifferentiated even in the absence of LIF, indicating that the presence of Prame17 can compensate for the need of MEK/ERK inhibition (PD0325901) (ref. 15). The considerably higher expression level of Prame17 in ICM compared with ESCs and its requirement to sustain pluripotency suggested an important role in naive pluripotency.

RESULTS

Expression of Prame17 in ESCs induces a gene signature similar to preimplantation naive epiblast cells

Analysis of Prame17 protein levels in blastocysts revealed specific expression in ICM cells and absence in trophectoderm[15,16] (Fig. 1a). Interestingly, *Prame17* expression was also higher in ESCs+2i than in ESCs+serum (Fig. 1b and Supplementary Fig. 1a). However, Prame17 levels in ESCs+2i were still very low and could not be detected by immunofluorescence, as done for ICM, or western blot with Prame17 antibodies (Fig. 1c and Supplementary Fig. 1a,b). The elevated Prame17 expression in ICM and its strong reduction during ICM–ESC transition[5] prompted us to investigate whether increasing the Prame17 expression might induce ESCs+serum to acquire a molecular signature similar to naive preimplantation epiblast cells. We made use of a previously established E14 ESC line that expresses elevated levels of FLAG-Prame17 (P7 ESCs) under the control of the

CAG promoter (Fig. 1c)[15]. We performed transcriptomic analysis and compared P7 ESCs versus E14 ESCs and found 1,190 genes displaying more than 1.5× higher transcript abundance in P7 ESCs ($P < 0.05$) whereas only 454 genes exhibited more than 1.5× higher expression in E14 ESCs (Fig. 1d and Supplementary Table 1). Transcription of the majority of categorized stem cell maintenance genes (Gene Ontology (GO): 0019827) including the core pluripotency factors *Oct4* and *Nanog* was not affected.

Since E14 ESCs and P7 ESCs are grown in serum, we analysed the RNA-seq results with published transcriptomic data comparing ICM or ESCs+2i versus ESCs+serum[5,12]. Of the 1,190 genes upregulated in P7 ESCs, 273 genes (23%) are upregulated in ICM versus ESCs (Fig. 1e). A similar number of ICM-specific genes are upregulated in ESCs+2i (352 out of 1,488 genes, 23%). In contrast, the number of genes downregulated was less in P7 ESCs (454) than in ESCs+2i (1,946), of which 124 and 650, respectively, were downregulated in ICM versus ESCs.

Functional annotation clustering of differentially expressed genes by GO revealed that P7 ESCs, ESCs+2i and ICM share similar biological processes (Supplementary Tables 2 and 3). The top 10 GO terms of genes with altered transcription are the same and all refer to developmental processes. Genes upregulated in P7 ESCs are highly enriched for terms associated with developmental process and the majority of them coincide with the top 10 GO terms of ICM-specific genes (Supplementary Table 4). In contrast, pathways associated with upregulated genes in ESCs+2i differ from ICM and P7 ESCs and, as previously reported, are highly enriched for terms associated with metabolic processes[12]. Downregulated genes in P7 ESCs are enriched for terms associated with ion transport and differ from GO terms associated with downregulated genes in ICM and ESCs+2i, which in turn are very similar to each other and implicated in developmental processes (Supplementary Table 5). Of note, *Fgf4*, a factor known to stimulate the ERK pathway that primes ESCs for lineage specification, is downregulated in P7 ESCs[17–19]. To assess how P7 ESCs are related to embryonic stages E3.5 and E4.5, and ESCs+serum or ESCs+2i, we performed principal component analysis (PCA) using recently published gene expression profiles[9]. Two-dimensional PCA analysis grouped ESCs+2i close to E4.5 epiblast cells (Fig. 1f)[9]. Remarkably, P7 ESCs were closer to ESCs+2i than to ESCs+serum and grouped with the greatest similarity to E4.5 epiblast cells. Dimensions 2 and 3 of the PCA analysis suggest that P7 ESCs also recapitulate certain aspects seen in epiblast cells (E3.5–4.5) (Supplementary Fig. 2).

Expression of Pramel7 induces hypomethylation of embryonic stem cells

The genome of ICM and ESCs+2i has a low content of DNA methylation compared with ESCs+serum and to later developmental stages[7,8,10,12]. To determine whether expression of Pramel7 affects the epigenetic state of ESCs+serum, we analysed the DNA methylation content of P7 ESCs. Intense 5-methylcytosine staining at 4′,6-diamidino-2-phenylindole (DAPI)-dense heterochromatic regions observed in control ESCs was significantly reduced in P7 ESCs (Fig. 2a). Measurement of global DNA methylation with the methylation-sensitive restriction enzyme HpaII indicated that CpG methylation was substantially reduced in P7 ESCs (Fig. 2b). We obtained similar results with McrBC digestion, an endonuclease that cleaves DNA containing only methylcytosine on one or both strands (Fig. 2c). Next, we deleted the *Pramel7* transgene with Cre-recombinase (P7^{-}) and found that HpaII or McrBC DNA digestion resistance reverted to levels found in control cells, indicating that the DNA methylation content depends on Pramel7 expression levels (Figs 1b and 2d).

Despite the low DNA methylation, ICM cells and ESCs+2i retain DNA methylation at imprinting control regions (ICRs)[8,10]. We examined the methylation of two ICRs, KvDMR1 and *Igf2/H19*, using bisulfite sequencing, COBRA and GlucMS-qPCR (glucosylation followed by methylation-sensitive quantitative PCR) and found that their methylation is drastically reduced in P7 ESCs (Fig. 2e,f and Supplementary Fig. 3a,b). This loss of methylation is also consistent with the upregulation of *H19* and *Kcnq1ot1* and silencing of the adjacent imprinted transcripts within the clusters (*Cdkn1c*) (Supplementary Table 1 and Supplementary Fig. 3c). A similar relationship between DNA methylation and gene expression was obtained through the analysis of methylation profiles of ESCs+serum[8], which revealed that genes upregulated following expression of Pramel7 display a higher methylation content at the transcription start site than average. As previously reported, genes upregulated following ESC transition from serum to 2i conditions do not show this correlation (Fig. 2g)[10]. These results indicate that expression of Pramel7 in ESCs+serum induces global hypomethylation and suggest a link between high Pramel7 expression levels and the hypomethylated state characterizing preimplantation embryos and ESCs+2i. Nevertheless, the methylation content of P7 ESCs is not identical to ICM and ESCs+2i since ICRs do not retain methylation.

Pramel7 interacts with UHRF1 and the Cullin 2 RING E3 ubiquitin ligase complex

P7 ESCs and E14 ESCs express similar levels of the DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b and the Ten-eleven translocation enzymes Tet1, Tet2 and Tet3 (Supplementary Fig. 4), indicating that Pramel7-mediated DNA hypomethylation is not due to altered expression of regulators of DNA methylation. To identify factors interacting with Pramel7, we performed a yeast two-hybrid screen using a Pramel7-LexA fusion protein as a bait to screen an E14 ESC complementary DNA library (Fig. 3a). Three out of five isolated clones displayed a strong LacZ signal, all expressing the protein UHRF1 (Ubiquitin-like, with PHD and RING finger domains 1). UHRF1 specifically recognizes and binds to hemimethylated DNA and is crucial for the transmission of DNA methylation marks during cell division[20–22]. Indeed, *Uhrf1*$^{-/-}$ ESCs display a drastic global reduction of DNA methylation content, including ICRs[20]. Anti-FLAG co-immunoprecipitation analysis in P7 ESCs confirmed the interaction of Pramel7 with UHRF1 (Fig. 3b). Pramel7–UHRF1 association was also detected in HEK293T cells, which do not express Pramel7, after anti-FLAG co-immunoprecipitation of cells transfected with plasmids expressing FLAG-Pramel7 (Fig. 3c,d). Additionally, we detected the association with histones, suggesting a role of Pramel7 that is linked to chromatin. We also identified Elongin C (TCEB1) and polyubiquitin as Pramel7-interacting proteins, both components of the Cullin 2 RING E3 ubiquitin ligase (CRL) complex that is involved in polyubiquitylation and subsequent proteasomal degradation of target substrates[23] (Fig. 3d). Interestingly, we found ubiquitin

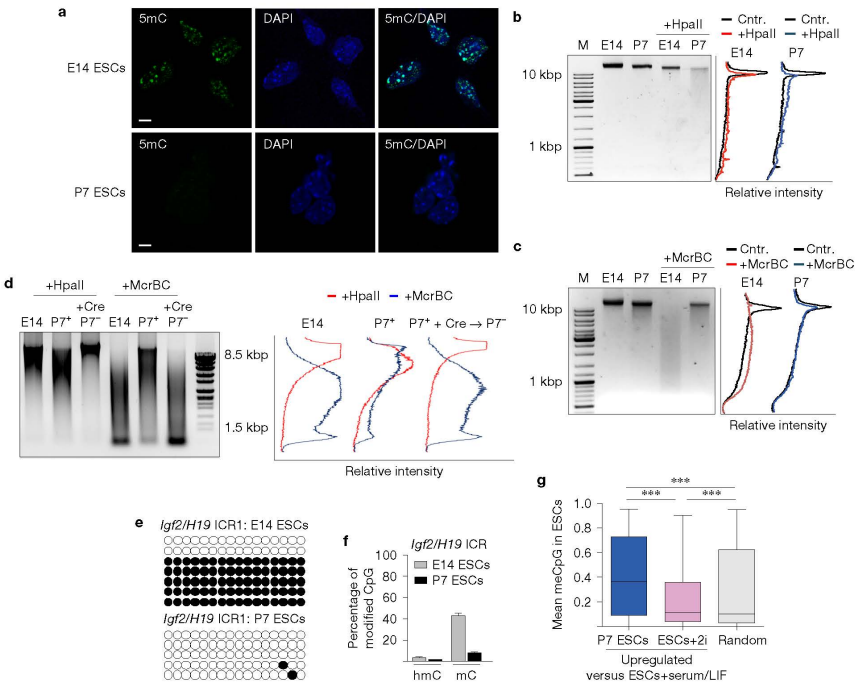


Figure 2 Expression of Pramel7 induces hypomethylation of embryonic stem cells. (**a**) DNA methylation defects in P7 ESCs in comparison with control E14 ESCs as revealed by immunofluorescent staining using anti-5mC antibody. Scale bars, 10 μm. The image is representative of three independent stainings. (**b**,**c**) CpG methylation levels measured by digestion of genomic DNA from E14 ESCs and P7 ESCs with the methylation-sensitive restriction enzymes HpaII and McrBC. Quantification of DNA signal was measured using Fiji image analysis software. M, DNA ladder. The image shown in **b** is representative of four independent experiments and in **c** of three independent experiments. (**d**) Defects in DNA methylation depend on Pramel7 expression levels. CpG methylation levels in E14, Cre-recombined (+Cre, P7$^-$) and non-recombined (P7$^+$) ESC clones were measured by HpaII and McrBC digestion. The image shown in **d** is representative of three independent experiments. (**e**) Bisulfite sequencing showing extensive demethylation of *Igf2/H19* ICR in P7 ESCs. (**f**) Analysis of methylation levels at one single MspI site (CCGG) by GlucMS-qPCR confirmed that *Igf2/H19* ICR underwent substantial demethylation without affecting hmC (hydroxy-methyl cytosine) levels in P7 ESCs. $n=3$ independent experiments performed in triplicates, error bars represent s.d. (**g**) CpG methylation levels at the transcription start site (+/− 1 kb) of RefSeq RNAs from differently expressed genes in P7 ESCs versus E14 ECSs and randomly picked (same number of transcripts as upregulated + downregulated) genes (data are from ref. 8). Box plot with median bar, first–third quantile box and 5th–95th percentile whiskers. (***$P<0.001$, two-tailed Student's *t*-test.)

associated with UHRF1 whereas the interaction with Elongin C was detected only in the presence of Pramel7 (Supplementary Table 6). Finally, analyses in *Uhrf1*$^{-/-}$ ESCs revealed that Pramel7 association with the CRL complex does not depend on UHRF1 (Fig. 3e,f).

Pramel7 mediates the proteasomal degradation of UHRF1

The above findings prompted us to investigate whether Pramel7 targets UHRF1 for proteasomal degradation. We measured UHRF1 levels by immunofluorescence and western blot analysis and found a significant reduction of UHRF1 in P7 ESCs (Fig. 4a,b). Similarly, UHRF1 levels were drastically reduced in HEK293T cells after 48 h of ectopic expression of Pramel7 (Fig. 4b). The reduction in UHRF1 was not mediated through the inhibition of *Uhrf1* gene expression, as *Uhrf1* messenger RNA levels in P7 ESCs were comparable to those of control ESCs (Fig. 4c). This is also consistent with a previous analysis showing no remarkable alterations of *Uhrf1* mRNA levels in ICM, where Pramel7 is highly expressed, and ESCs[5]. Remarkably, UHRF1 levels were also reduced in ESCs+2i when compared with ESCs+serum, a result that is consistent with a recent report[24] (Supplementary Fig. 1b).

To get insight into the Pramel7-mediated degradation of UHRF1, we excised the *Pramel7* cassette in P7 ESCs by Cre-recombinase, simultaneously bringing *EGFP* under the control of the CAG promoter (Fig. 1b). Clones were selected for EGFP-positive (P7$^-$/EGFP$^+$) and -negative (P7$^+$/EGFP$^-$) signal. The dependency of UHRF1 moieties on Pramel7 expression was evident by the reduction of UHRF1

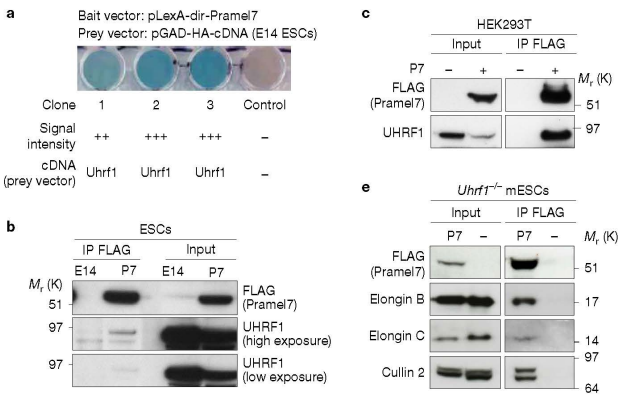


Figure 3 Pramel7 associates with UHRF1 and components of the Cullin 2 RING E3 ubiquitin ligase (CRL) complex. (**a**) Yeast two-hybrid assay using Pramel7 as a bait and a whole transcriptome cDNA library of E14 ESCs as prey. Three clones displayed a strong LacZ signal and all expressed UHRF1. (**b**,**c**) FLAG-immunoprecipitation from P7 ESCs and HEK293T cells transfected with FLAG-Pramel7 (P7)-expressing plasmids. Immunoblots show association of Pramel7 with UHRF1. Images are representative of five independent experiments. (**d**) Mass-spectrometry analysis of proteins co-immunoprecipitated with FLAG-Pramel7 (P7) in HEK293T cells. Values represent peptide hits. (**e**) FLAG co-immunoprecipitation of components of the Cullin 2 RING E3 ubiquitin ligase (CRL) complex (Elongin B, Elongin C, Cullin 2) in *Uhrf1*$^{-/-}$ ESCs expressing FLAG-Pramel7 (P7). The image is representative of three independent experiments. mESC, mouse ESC. (**f**) Mass-spectrometry analysis of proteins FLAG-immunoprecipitated in *Uhrf1*$^{-/-}$ ESCs expressing FLAG-Pramel7 (P7). Values represent peptide hits. Unprocessed original scans of immunoblots are shown in Supplementary Fig. 6.

expression in P7^{+}/EGFP^{-} cells as measured by immunofluorescence and western blotting (Fig. 4d,e). Finally, we monitored UHRF1 levels in HEK293T cells after transfection of FLAG-Pramel7 expression plasmid (Fig. 4f). Pramel7 expression was first detected 24 h post-transfection, which coincides with the initial UHRF1 reduction. After 72 h, when Pramel7 expression started to decline, UHRF1 again increased and, concomitant with the disappearance of Pramel7 (96 h and 120 h), UHRF1 levels were restored to similar amounts found in control cells.

The 26S-proteasome mediates the final step in the degradation of polyubiquitin-tagged proteins. We evaluated endogenous ubiquitylation of UHRF1 and found that ectopic Pramel7 expression increases polyubiquitylated UHRF1 (Fig. 4g). Next, we analysed whether UHRF1 is modified with Lys48-linked polyubiquitin chains, the canonical signal for proteasomal degradation[25]. We co-expressed UHRF1 together with the FLAG-tagged ubiquitin wild-type (WT) and the mutants Lys0, which can only monoubiquitylate target proteins, and Lys48-only, which can generate only Lys48-ubiquitin chains. We found that UHRF1 polyubiquitylation is largely due to Lys48-linkage of ubiquitin (Fig. 4h) and that the presence of Pramel7 further increased UHRF1 Lys48-polyubiquitylation. Finally, we tested the effect of proteasomal inhibition on the stability of UHRF1. Treatment with the proteasome inhibitor MG132 increases UHRF1 amounts and blocked Pramel7-mediated UHRF1 degradation (Fig. 4i), indicating that Pramel7 targets UHRF1 for degradation via the 26S-proteasome pathway. Thus, genome hypomethylation observed in P7 ESCs correlates well with the degradation of UHRF1 (ref. 20). Furthermore, the data suggest that the reduction of DNA methylation following Pramel7 expression occurs via a passive demethylation process, which includes the loss of methylation at ICRs as previously reported in the case of *Uhrf1*$^{-/-}$ ESCs[20].

The Pramel7-LRR and UHRF1-SRA and RING domains are necessary for the Pramel7–UHRF1 interaction and UHRF1 degradation

We next analysed whether the interaction of Pramel7 with UHRF1 is required for UHRF1 degradation. Pramel7 contains 3 leucine-rich repeat (LRR) motifs, which in general are implicated in the formation of protein–protein interactions[26]. Analysis of Pramel7-LRR deletion mutants indicated that deletion of the first LRR (ΔLRR1) was sufficient to impair Pramel7-mediated degradation of UHRF1 and the association of Pramel7 with UHRF1 (Fig. 5a–c). Similar results were obtained with further deletions of the second (ΔLRR1/2) and third LRR (ΔLRR1/2/3). Thus, Pramel7-mediated degradation of UHRF1 requires the association of UHRF1 with the Pramel7-LRR motif. Next, we analysed the degradation and Pramel7 association with UHRF1 mutants lacking the carboxy-terminal region, including all PHD, SRA and RING (ΔPSR) domains or only SRA and RING motifs (ΔSR) (Fig. 5d–f). These regions were previously implicated in hemimethylated DNA recognition, ubiquitylation of histone H3K23 and association with the histone methyltransferase G9a (refs 27–29). Deletion of SRA and RING regions was sufficient to impair UHRF1 degradation and Pramel7 interaction with both ΔPSR and ΔSR UHRF1 mutants was strongly reduced (Fig. 5e,f). UHRF1 associates with histones and, consistent with previous reports[30], this interaction is abolished following deletion of PSR domains, whereas the ΔSR mutant still associates with histones (Fig. 5g). Similar to UHRF1, Pramel7 in ESCs also associated with histones (Figs 3d and 5h) and

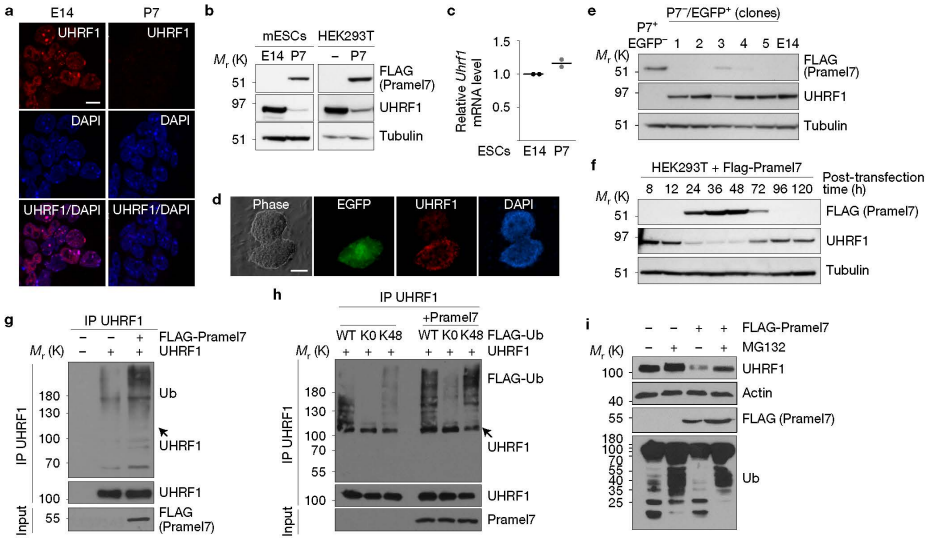


Figure 4 Pramel7 expression affects UHRF1 stability and leads to UHRF1 degradation via the 26S-proteasome pathway. (**a**) Immunofluorescent labelling of UHRF1 protein in E14 ESCs and P7 ESCs (P7). Scale bar, 10 μm. (**b**) Western blot for Pramel7 and UHRF1 in whole-cell lysates of E14 ESCs and P7 ESCs and in HEK293T cells 48 h after transient transfection with FLAG-Pramel7 plasmid (P7). Tubulin is shown as a loading control. (**c**) *Uhrf1* gene expression in E14 ESCs and P7 ESCs. *Uhrf1* mRNA levels were measured by qRT–PCR and normalized to actin mRNA. Average values of two independent experiments; lines represent means. (**d**) Immunofluorescence showing mutually exclusive UHRF1 and EGFP expression in P7$^-$/EGFP$^+$ (Cre-recombined) and P7$^+$/EGFP$^-$ (non-recombined) ESC colonies; scale bar, 200 μm. (**e**) UHRF1 and Pramel7 levels measured by western blotting analysis of cell lysates from five P7$^-$/EGFP$^+$ ESC clones and one P7$^+$/EGFP$^-$ ESC clone. (**f**) UHRF1 levels in HEK293T cells were monitored during the 5 days after transient transfection of FLAG-Pramel7 plasmid DNA (8 h to 120 h post-transfection). (**g**) Pramel7 increases polyubiquitylation of UHRF1. Endogenous ubiquitin levels of UHRF were monitored following immunoprecipitation of UHRF1 from HEK293T cells transfected with plasmid expressing UHRF1 and/or FLAG-Pramel7. (**h**) UHRF1 is modified through Lys48-linked polyubiquitin chains. Ubiquitin levels of UHRF1 were measured following immunoprecipitation of UHRF1 from HEK293T cells transfected with plasmid expressing UHRF1 and WT, Lys0 (K0) or Lys48 (K48) Flag-ubiquitin as indicated. (**i**) Pramel7-mediated UHRF1 degradation in HEK293T cells was monitored after addition of the proteasome inhibitor MG132 (20 μM). Images in **a**,**b**,**d**–**i** are representative of three independent experiments. Unprocessed original scans of immunoblots are shown in Supplementary Fig. 6.

exhibited a preferential association with chromatin, which as expected was also enriched in UHRF1 (Fig. 5i). We conclude that Pramel7-LRR and UHRF1-SRA and RING domains mediate Pramel7–UHRF1 interaction, which is required to target UHRF1 to the 26S-proteasome. Furthermore, this association is most likely occurring on chromatin.

Lack of Pramel7 increases UHRF1 protein levels in embryos and impairs blastocyst formation

Our attempts to obtain Pramel7-null mice and blastocysts have failed so far, suggesting an important role of Pramel7 in the late stages of preimplantation development. To better understand the *in vivo* function of Pramel7 we made use of the CRISPR/Cas9 technology to target the coding region of Pramel7 by using a single guide RNA (sgRNA-P7). As a control, we used sgRNA targeting the *Rosa26* locus[31]. sgRNA-P7 and sgRNA-Rosa26 were microinjected into the pronuclei of C57Bl/6J zygotes. The surviving zygotes (E0.5) were cultured for 4 days (until E4.5) and the development of the embryos was checked daily. Embryos developed normally until E3.5 (Fig. 6a). At E4.5 the number of sgRNA-P7 embryos that reached the blastocyst stage drastically reduced compared with the controls and was mirrored in an increase of developmentally arrested morula. Arrested sgRNA-P7 embryos showed an aberrant morphology and lacked expression of Pramel7 (Pramel7-KO) (Fig. 6b,c). sgRNA-P7 embryos that reached blastocyst stage were still expressing Pramel7, indicating that only a successful Pramel7 deletion causes a developmental arrest and that Pramel7 is required for the establishment of the blastocyst. Correctly developed embryos displayed a mutually exclusive expression of Pramel7 and UHRF1 in ICM whereas Pramel7-KO embryos expressed high UHRF1 (Fig. 6c). These results support a role of Pramel7 in regulating UHRF1 stability in ICM.

Pramel7 maintains the pluripotency state by repressing DNA methylation through regulation of UHRF1 stability

Following differentiation, ESC chromatin remodels into a condensed and repressed structure and DNA methylation was shown to be crucial for permanent restriction of developmental fate during

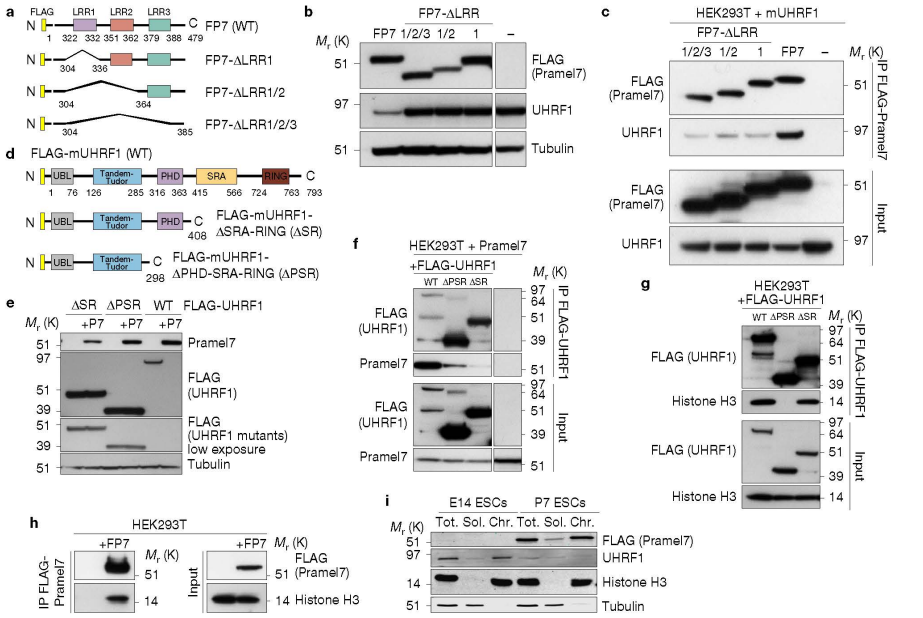


Figure 5 Pramel7 LRRs and UHRF1-SRA and RING domains are implicated in Pramel7–UHRF1 interaction and UHRF1 stability. (**a**) Schema representing the domain organization of Pramel7 and the analysed ΔLRR mutants. LRR, leucine-rich repeat. (**b**) Western blot of UHRF1 in HEK293T cells 48 h after transfection of FLAG-Pramel7 (FP7) and FP7-ΔLRR mutants; tubulin is shown as a loading control. (**c**) FLAG immunoprecipitation from HEK293T cells expressing FP7-ΔLRR mutants and mouse UHRF1 (mUHRF1). Overexpression of mUHRF1 was intended to increase the UHRF1 signal and achieve better detection of the UHRF1–Pramel7 interaction. (**d**) Schematic representation of FLAG-UHRF1 and UHRF1 mutant constructs: ΔSR (deletion of SRA and RING domains) and ΔPSR (deletion of PHD, SRA and RING domains). (**e**) Western blot of UHRF1 WT and mutants in HEK293T cells 48 h post-transfection of UHRF1 and Pramel7 expression plasmids. (**f**) FLAG-immunoprecipitation from HEK293T cells transfected with FLAG-UHRF1 WT and mutants and Pramel7 expression plasmids. (**g**,**h**) Association of UHRF1 and Pramel7 with histone H3 shown by FLAG-immunoprecipitation from HEK293T expressing FLAG-UHRF1 WT or FLAG-UHRF1 mutants (**g**) and FLAG-Pramel7 (**h**). (**i**) UHRF1 and Pramel7 associate with chromatin of ESCs. Chromatin-bound (Chr.) and soluble (Sol.) fractions of equivalent cell number of E14 ESCs and P7 ESCs were analysed by western blot for UHRF1 and Pramel7 levels. Tot., total. Tubulin and histones are shown as loading and fractionation control. Images in **b**,**c**,**e**–**h** are representative of three independent experiments. Unprocessed original scans of immunoblots are shown in Supplementary Fig. 6.

differentiation[32–34]. To determine whether the expression of Pramel7 following ESC differentiation impairs terminal differentiation, we tested the ability of P7 ESCs to differentiate into a stable state. We differentiated P7 ESCs following withdrawal of LIF and feeders for 14 days and observed that E14 ESCs developed into a homogeneous cell layer whereas P7 ESCs formed less differentiated and more compact colonies (Fig. 7a). As expected, UHRF1 levels were lower in differentiated cells than in ESCs reflecting a lower cell proliferation rate of the differentiated state[35] (Fig. 7b). Pramel7-mediated UHRF1 degradation was still active during differentiation as evident by low UHRF1 and DNA methylation content (Fig. 7b,c). In differentiated P7 cells the pluripotency genes *Oct4*, *Nanog* and *Rex1* and the pluripotency marker stage-specific embryonic antigen 1 (SSEA1) were expressed at high levels, suggesting that transcriptional silencing was not as efficient as in differentiated control ESCs (Fig. 7a,d). Consistent with the hypomethylated state of P7 ESC genome, methylation of the *Oct4* promoter was lower in differentiated P7 cells than in control-differentiated cells (Fig. 7e). *Uhrf1*$^{-/-}$ ESCs shared several similarities with P7 ESCs. After 26 days of differentiation, *Uhrf1*$^{-/-}$ ESCs were also defective in transcriptional silencing of pluripotency genes and formed less differentiated and more compact colonies, a phenotype that we also observed in DNMT-TKO cells (Fig. 7f and Supplementary Fig. 5a). Finally, to determine whether expression of Pramel7 affects the stable terminally differentiated state, we tested the ability of 14-day differentiated P7 cells to reconvert to a pluripotent state (Fig. 8a). Differentiated cells were seeded at a density of 10 cells per well in 96-well plates, cultured for 7 days in conditions supporting pluripotency (+LIF) and screened for the expression of the pluripotency marker alkaline phosphatase (AP). AP-positive (AP^{+}) colonies were found in only 2.7% of E14 ESC wells (5/182) whereas 80.8% (147/182) of wells with P7 ESCs contained AP^{+} colonies. P7 ESC reconverted clones maintained Pramel7 expression and low

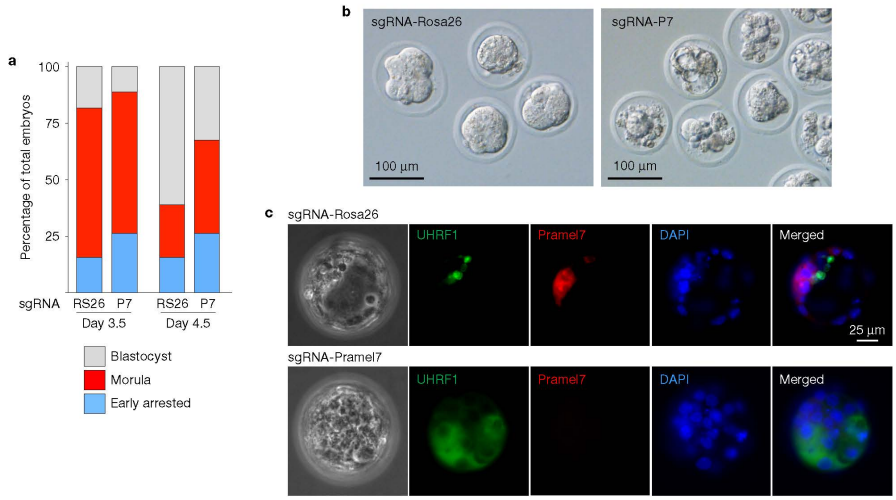


Figure 6 Pramel7 is required for ICM. (**a**) Stacked bar plots showing the fraction of embryos at different developmental stages at day 3.5 and day 4.5, after injection with sgRNA-Rosa26 and sgRNA-P7. Data are from two independent microinjections. In experiment 1, 114 embryos were microinjected with sgRNA-Rosa26 and 110 embryos with sgRNA-P7. In experiment 2, 87 embryos were microinjected with sgRNA-Rosa26 and 91 embryos with sgRNA-P7. Source data for this experiment are available in Supplementary Table 8. (**b**) Representative images of embryos at the morula stage obtained at day 4.5 showing aberrant morphology of sgRNA-P7-injected embryos compared with sgRNA-Rosa26 control embryos. (**c**) Immunofluorescence analysis of a blastocyst from a sgRNA-Rosa26 control embryo showing mutually exclusive expression of UHRF1 and Pramel7 in ICM cells. sgRNA-Pramel7 embryos lacked Pramel7 expression and expressed UHRF1.

UHRF1 levels (Fig. 8b) and expressed pluripotency markers (Fig. 8c and Supplementary Fig. 5b). We conclude that increased expression of Pramel7 in ESCs impairs exit from pluripotency by repressing DNA methylation through regulation of UHRF1 stability.

DISCUSSION

ESCs need to be shielded from extrinsic signals to indefinitely retain pluripotency in culture and serum/LIF and 2i conditions have both been proposed to be conducive for maintenance of naive pluripotency. Additionally, 2i optimizes the state of naive pluripotency, inducing a gene signature and DNA hypomethylation state that resembles naive epiblast cells[36]. Here we show that expression levels of Pramel7 correlate with ground-state pluripotency, high in ICM and ESCs+2i and low in ESCs+serum (Fig. 8d). Increasing Pramel7 expression in ESCs+serum causes global DNA hypomethylation and induces a naive pluripotency state closely comparable to the developmental ground state *in vivo*.

ESCs+2i or P7 ESCs have similar properties, for example, robust pluripotency, hypomethylated genome and activated pathways, all similar to ICM. Nevertheless, it is likely that they achieved these features using different mechanisms. Genes upregulated in P7 ESCs are strongly linked to developmental processes and coincide with ICM-specific genes whereas genes downregulated in ESCs+2i share more similarity with ICM. Additionally, P7 ESCs are characterized by elevated expression of genes involved in sustaining self-renewal and retention of pluripotency such as *Gbx2* and *Lef1*[37,38] and reduced expression of genes implicated in controlling exit from pluripotency such as *Fgf4* and *Tbx3*[17–19,39]. Therefore, an inducible system that modulates expression of Pramel7 in cultured pluripotent ESCs might represent an attractive and physiological alternative to *in vitro* reproduce the *in vivo* ground-state pluripotency.

DNA hypomethylation has been considered an epigenetic trait of ground-state pluripotency. Downregulation of both the *de novo* DNA methyltransferases Dnmt3a and Dnmt3b has been proposed as a way to maintain genome hypomethylation in ICM[11]. A similar mechanism has been observed in ESCs+2i (refs 8,11). The post-translational control of UHRF1 levels through Pramel7 might represent an additional regulatory pathway to establish a hypomethylated genome, creating greater epigenetic flexibility required by preimplantation epiblasts to undergo rapid and transient developmental changes. Interestingly, during the revision of this work, it was reported that ESCs+2i have lower UHRF1 protein levels compared with ESCs+serum[24]. Our data support these results. Moreover, we showed that Pramel7 levels are higher in ESCs+2i than in ESCs+serum. Although the role of Pramel7 in ESCs+2i still requires further investigation, these results are consistent with our *in vivo* analyses determining a mutually exclusive expression of Pramel7 and UHRF1 in ICM cells. The phenotype of P7 ESCs and *Uhrf1*$^{-/-}$ ESCs is very similar, including the loss of methylation at ICRs and their corresponding transcriptional changes[20]. However, ICM and ESCs+2i

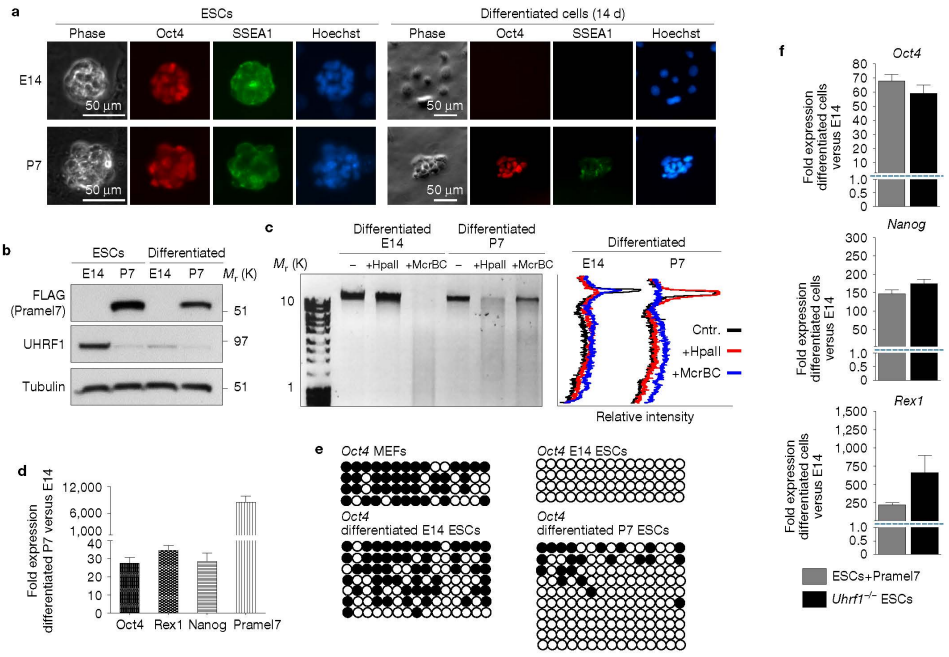


Figure 7 Pramel7 maintains the pluripotent state by repressing DNA methylation through regulation of UHRF1 stability. (**a**) Immunofluorescence analysis of the pluripotency markers Oct4 and SSEA1 in E14 ESCs and P7 ESCs and after 14 days of differentiation. Images are representative of four independent stainings. (**b**) UHRF1 and Pramel7 protein levels from whole-cell lysates of E14 ESCs and P7 ESCs in self-renewing conditions (+LIF) and after 14 days of differentiation. Representative of three independent experiments. (**c**) DNA methylation levels of 14-day differentiated E14 ESCs and P7 ESCs measured by digestion with HpaII and McrBC. Representative of three independent experiments. (**d**) Expression of the pluripotency-associated genes *Oct4*, *Nanog* and *Rex1* in differentiated cells measured by qRT–PCR and normalized to actin mRNA. Data are represented as fold expression in differentiated P7 cells compared with differentiated E14 ESCs. Values represent the average of $n = 3$ independent experiments performed in triplicates; error bars represent s.d. (**e**) Bisulfite genomic sequencing analysis of the *Oct4* promoter in MEFs, E14 ESCs and differentiated E14 and P7 cells. Open and filled circles indicate unmethylated and methylated CpGs, respectively. (**f**) *Oct4*, *Rex1* and *Nanog* expression levels in P7 ESCs and *Uhrf1*$^{-/-}$ ESCs after 26 days of differentiation. Values were measured by qRT–PCR, normalized to actin mRNA and represented as fold change relative to the expression of differentiated E14 ESCs. Values represent the average of $n = 3$ independent experiments performed in triplicates, error bars represent s.d. Unprocessed original scans of immunoblots are shown in Supplementary Fig. 6.

maintain the parental methylation pattern at imprinted loci[7,10,11]. There are several reasons that might explain the loss of methylation at ICRs in P7 ESCs. Compared with ICM, which undergoes only a few rounds of duplication, P7 ESCs have divided for a much longer time and thus are more prone to lose methylation, including methylation at ICRs (as in the case of *Uhrf1*$^{-/-}$ ESCs[20]). Moreover, UHRF1 in ESCs+2i is only modestly reduced while the elevated expression of Pramel7 in P7 ESCs (under the control of CAG promoter) induced a strong UHRF1 downregulation. Alternatively, ICM cells might modulate the activity of Pramel7 at defined gene loci (that is, at ICRs) whereas the elevated expression of Pramel7 in P7 ESCs might not allow this tight regulation.

Our results suggest that controlling the maintenance of DNA methylation through modulation of UHRF1 stability has the potential to regulate cell fate and might represent an as-yet-unappreciated dynamic nature of DNA methylation. Our data imply that maintenance of DNA methylation has a role for stable exit from pluripotency, as evident by differentiation defects observed in ESCs with deletion of *Dnmt1* (ref. 32) or *Uhrf1* or Pramel7-mediated degradation of UHRF1. Importantly, demethylation of ESCs following knockout of all *Dnmt* or *Uhrf1* genes or following culture in 2i conditions did not result in a proper activation of methylated genes[10,40,41], which suggests that either transcription factors are not present or compensatory mechanisms can replace the repressive effect of DNA methylation. The fact that increased expression of Pramel7 induced global demethylation and changes in transcription at genes with high methylation content indicates that Pramel7 might have additional functions not limited to DNA methylation pathways.

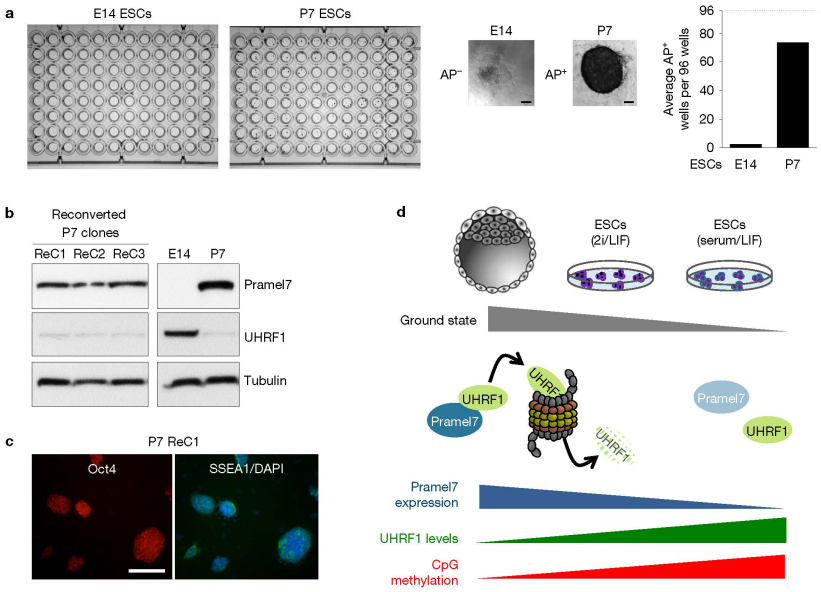


Figure 8 Forced expression of Pramel7 affects the stable terminally differentiated state. (**a**) Alkaline phosphatase (AP) staining assay of 14-day differentiated E14 ESCs and P7 ESCs re-exposed to self-renewing conditions (+LIF) for 7 days. Representative images of AP-stained 96-well plates and of an AP⁻ colony typical for E14 ESCs and an AP⁺ colony characteristic for P7 ESCs are shown. Scale bars, 100 μm. The right panel shows average numbers of wells containing AP⁺ colonies. The experiment was repeated $n=2$ independent times. (**b**) Differentiated P7 cells reconverted into ESCs (ReCs) still exhibit decreased levels of UHRF1. Western blot indicated UHRF1 protein levels of three P7 reconverted clones expanded in self-renewing conditions (+LIF). Western blot representative of two independent experiments. (**c**) Representative immunofluorescence of a P7 reconverted clone (P7-ReC1) showing expression of the pluripotency markers Oct4 and SSEA1. Scale bar, 200 μm. The image is representative of three independent experiments. (**d**) Model showing how Pramel7 contributes to ground-state pluripotency. Expression levels of Pramel7 correlate with ground-state pluripotency, high in ICM and ESCs+2i and low in ESCs in serum/LIF. Pramel7 induces global hypomethylation by targeting UHRF1 for proteasomal degradation and establishes an active epigenetic state that might be required by ICM to undergo rapid and transient developmental changes. Unprocessed original scans of immunoblots are shown in Supplementary Fig. 6.

The post-transcriptional control of UHRF1 levels mediated by Pramel7 joins two pathways, proteasome and epigenetics, which have been independently implicated in the maintenance of pluripotency. We showed that controlling UHRF1 levels through Pramel7-mediated targeting for proteasome degradation is an important molecular pathway for the regulation of epigenetic states linked to ground-state pluripotency (Fig. 8d). Since PRAME and PRAME family members are expressed in a wide range of human tumours and are often correlated with poor clinical outcome[42–44], our data also offer important insights into how developmental programs might be undermined, leading to the formation of diseased tissues, including cancers. □

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of this paper.

Note: Supplementary Information is available in the online version of the paper

ACKNOWLEDGEMENTS

The authors thank D. Bär for technical assistance, W. Piwko and O. Shakhova for discussions and reagents and J. vom Berg, D. Korkmaz and M. Tarnowska for embryo isolation. We acknowledge assistance provided by the Functional Genomic Center Zurich, especially C. Aquino and L. Opitz. This work was supported by the Olga Mayenfisch Foundation (to P.C. and R.S.), Julius Müller Stiftung (to R.S.), the Novartis Foundation for Medical-Biological Research (to P.C.), the Theodor und Ida Herzog-Egli Foundation (to P.C. and R.S.), the Sassella Stiftung (P.C.), the Stiftung für Wissenschaftliche Forschung an der Universität Zürich (to P.C. and R.S.), the Helmut Horten Stiftung (to L.P.), Krebsliga Schweiz (KFS-3497-08-2014 to R.S.), the Swiss National Science Foundation (31003A_173056 and 31003A-152854 to R.S., 323530-133905 to F.A.W., 31003A-166370 to L.P.), the UBS-Promedica Stiftung (to R.S.) and the Forschungskredit of the University of Zurich (to U.G., E.V. and D.D.).

AUTHOR CONTRIBUTIONS

U.G. performed most of the experiments, contributed to experimental design and data analysis and wrote the manuscript. E.A.C. performed yeast two-hybrid screening. S.W., D.D. and E.V. performed methylation analysis. R.S., D.D. and M.W.S. analysed transcriptome and methylation data. M.J.O. performed RNA-seq and PCA analysis. E.A.C., F.A.W. and S.S.P. generated and analysed Pramel7 ESCs. J.L., J.S., H.K. and J.W. contributed to UHRF1 plasmids and UHRF1-KO ESCs. G.A.W. contributed to data analysis. M.G. and L.P. performed the ubiquitylation analysis. P.C. and P.P. generated and analysed Pramel7 knockout embryos; P.C. designed most of the experiments; R.S. and P.C. jointly directed the study and wrote the manuscript.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Published online at http://dx.doi.org/10.1038/ncb3554

1. Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl Acad. Sci. USA* **78,** 7634–7638 (1981).
2. Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292,** 154–156 (1981).
3. Zwaka, T. P. & Thomson, J. A. A germ cell origin of embryonic stem cells? *Development* **132,** 227–233 (2005).
4. Niwa, H. How is pluripotency determined and maintained? *Development* **134,** 635–646 (2007).
5. Tang, F. *et al.* Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell* **6,** 468–478 (2010).
6. Hackett, J. A. & Surani, M. A. Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* **15,** 416–430 (2014).
7. Smith, Z. D. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484,** 339–344 (2012).
8. Habibi, E. *et al.* Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **13,** 360–369 (2013).
9. Boroviak, T., Loos, R., Bertone, P., Smith, A. & Nichols, J. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat. Cell Biol.* **16,** 516–528 (2014).
10. Ficz, G. *et al.* FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* **13,** 351–359 (2013).
11. Leitch, H. G. *et al.* Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20,** 311–316 (2013).
12. Marks, H. *et al.* The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149,** 590–604 (2012).
13. Ying, Q. L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453,** 519–523 (2008).
14. Bortvin, A. *et al.* Incomplete reactivation of Oct4-related genes in mouse embryos cloned from somatic nuclei. *Development* **130,** 1673–1680 (2003).
15. Casanova, E. *et al.* Pramel7 mediates LIF/STAT3-dependent self-renewal in embryonic stem cells. *Stem Cells* **29,** 474–485 (2011).
16. Cinelli, P. *et al.* Expression profiling in transgenic FVB/N embryonic stem cells overexpressing STAT3. *BMC Dev. Biol.* **8,** 57 (2008).
17. Burdon, T., Stracey, C., Chambers, I., Nichols, J. & Smith, A. Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Dev. Biol.* **210,** 30–43 (1999).
18. Kunath, T. *et al.* FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134,** 2895–2902 (2007).
19. Stavridis, M. P., Lunn, J. S., Collins, B. J. & Storey, K. G. A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development* **134,** 2889–2894 (2007).
20. Sharif, J. *et al.* The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450,** 908–912 (2007).
21. Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317,** 1760–1764 (2007).
22. Liu, X. *et al.* UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nat. Commun.* **4,** 1563 (2013).
23. Hotton, S. K. & Callis, J. Regulation of cullin RING ligases. *Annu. Rev. Plant Biol.* **59,** 467–489 (2008).
24. von Meyenn, F. *et al.* Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol. Cell* **62,** 848–861 (2016).
25. Grice, G. L. & Nathan, J. A. The recognition of ubiquitinated proteins by the proteasome. *Cell. Mol. Life Sci.* **73,** 3497–3506 (2016).
26. Kobe, B. & Kajava, A. V. The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol.* **11,** 725–732 (2001).
27. Nishiyama, A. *et al.* Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* **502,** 249–253 (2013).
28. Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. & Shirakawa, M. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* **455,** 818–821 (2008).
29. Achour, M. *et al.* UHRF1 recruits the histone acetyltransferase Tip60 and controls its expression and activity. *Biochem. Biophys. Res. Commun.* **390,** 523–528 (2009).
30. Gelato, K. A. *et al.* Accessibility of different histone H3-binding domains of UHRF1 is allosterically regulated by phosphatidylinositol 5-phosphate. *Mol. Cell* **54,** 905–919 (2014).
31. Chu, V. T. *et al.* Efficient generation of Rosa26 knock-in mice using CRISPR/Cas9 in C57BL/6 zygotes. *BMC Biotechnol.* **16,** 4 (2016).
32. Schmidt, C. S. *et al.* Global DNA hypomethylation prevents consolidation of differentiation programs and allows reversion to the embryonic stem cell state. *PLoS ONE* **7,** e52629 (2012).
33. Sakaue, M. *et al.* DNA methylation is dispensable for the growth and survival of the extraembryonic lineages. *Curr. Biol.* **20,** 1452–1457 (2010).
34. Savic, N. *et al.* lncRNA maturation to initiate heterochromatin formation in the nucleolus is required for exit from pluripotency in ESCs. *Cell Stem Cell* **15,** 720–734 (2014).
35. Bonapace, I. M. *et al.* Np95 is regulated by E1A during mitotic reactivation of terminally differentiated cells and is essential for S phase entry. *J. Cell Biol.* **157,** 909–914 (2002).
36. Kalkan, T. & Smith, A. Mapping the route from naive pluripotency to lineage specification. *Phil. Trans. R Soc. B* **369,** 20130540 (2014).
37. Tai, C. I. & Ying, Q. L. Gbx2, a LIF/Stat3 target, promotes reprogramming to and retention of the pluripotent ground state. *J. Cell Sci.* **126,** 1093–1098 (2013).
38. Huang, C. & Qin, D. Role of Lef1 in sustaining self-renewal in mouse embryonic stem cells. *J. Genet. Genomics* **37,** 441–449 (2010).
39. Weidgang, C. E. *et al.* TBX3 directs cell-fate decision toward mesendoderm. *Stem Cell Rep.* **1,** 248–265 (2013).
40. Fouse, S. D. *et al.* Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* **2,** 160–169 (2008).
41. Sharif, J. *et al.* Activation of endogenous retroviruses in Dnmt1$^{-/-}$ ESCs involves disruption of SETDB1-mediated repression by NP95 binding to hemimethylated DNA. *Cell Stem Cell* **19,** 81–94 (2016).
42. van Baren, N. *et al.* PRAME, a gene encoding an antigen recognized on a human melanoma by cytolytic T cells, is expressed in acute leukaemia cells. *Br. J. Haematol.* **102,** 1376–1379 (1998).
43. Ikeda, H. *et al.* Characterization of an antigen that is recognized on a melanoma showing partial HLA loss by CTL expressing an NK inhibitory receptor. *Immunity* **6,** 199–208 (1997).
44. van 't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415,** 530–536 (2002).

DOI: 10.1038/ncb3554

METHODS

Animals. All experiments were performed with C57BL/6J mice. Fertilized eggs were obtained from superovulated 3–4-week-old females. Animals were housed under controlled lighting (lights on at 6:00–18:00), temperature (23 ± 2 °C) and humidity (50 ± 5%), with free access to food and water. The Veterinary Offices of the Canton of Zurich, Switzerland (license no. 165/2014 to P.C.) and Basel, Switzerland (license no. 1023G1 to P.P.) approved all animal experiments. Housing and experimental procedures were in accordance with the Swiss animal protection law and conformed to the European Convention for the protection of vertebrate animals used for experimental and other scientific purposes (Council of Europe no. 123, Strasbourg 1985).

Cell lines and culture media. No cell lines used in this study were found in the database of commonly misidentified cell lines that is maintained by ICLAC and NCBI Biosample. All of the cell lines generated for this study were derived from E14 129/Ola, karyotyped, tested for mycoplasma and their identity authenticated by several means including resistance to drug selection, PCR and sequencing of specific genomic regions.

ESCs were routinely cultivated on mitotically inactivated mouse embryonic fibroblasts (MEFs) in complete medium (Serum +LIF: GMEM (Sigma), 10% FCS, 10 mM sodium pyruvate, 1× NEAA, 1× Pen/Strep/Glu, 0.1 mM 2-mercaptoethanol) supplemented with 1,000 U ml^{-1} leukaemia inhibitory factor (LIF). HEK293T cells and MEFs were grown in DMEM (Life Technologies) supplemented with 10% FCS, 1× Pen/Strep/Glu and 10 mM sodium pyruvate. Establishment of Pramel7 ESCs (P7 ESC) was described in ref. 15. *Uhrf1*$^{-/-}$ ESCs were previously described[20]. DNMT-TKO ESCs were kindly provided by M. Okano[45] (Laboratory for Mammalian Epigenetic Studies, Center for Developmental Biology, RIKEN, Kobe, Japan).

Immunofluorescent detection and alkaline phosphatase staining. Cells were fixed in 4% formalin. For confocal analysis formalin was supplemented with 1:555 100% Triton X-100 (Sigma). Primary antibodies were diluted in PBST (PBS + 0.1% Tween-20) and 4% horse serum and cells were incubated either overnight at 4 °C or at room temperature for 2–4 h. Nuclei were stained with DAPI (Roche) or Hoechst for 60 s. Wide-field images were taken on a Zeiss Axiovert 40 CFL and processed using AxioVision 4.6 software (Zeiss) and Adobe Photoshop CS6. Confocal pictures were taken in the Center for Microscopy and Image Analysis (ZMB) of the University of Zurich and processed with IMARIS 7.6 software. For alkaline phosphatase staining, cells were fixed in 4% formalin, washed 2× with AP buffer (0.1 M Tris-HCl, 0.1 M NaCl, 20 mM $MgCl_2$, pH 9.5) and incubated with AP staining solution (AP Buffer +0.5 µl ml^{-1} NBT (Roche) and 3.5 µl ml^{-1} BCIP (Roche)). The staining reaction was stopped with 1× Tris-EDTA (20 mM Tris-HCl, 5 mM EDTA).

Isolation of genomic DNA and DNA methylation analysis. Cells were incubated with Proteinase K overnight at 50 °C followed by treatment with RNAseA (Fermentas) for 30 min at 37 °C. Genomic DNA (gDNA) was purified with 25:24:1 phenol/chloroform/isoamyl alcohol and subsequently precipitated by addition of NH_4Ac/ethanol. Four micrograms of gDNA was digested overnight with 20 U HpaII or McrBC in 20 µl total reaction volume containing 0.5 ng of pBluescript plasmid DNA for testing the efficiency of the HpaII digest. Digested gDNA was loaded on a 0.8% agarose gel and gDNA fragments were separated by electrophoresis. Quantification of DNA signal was measured using Fiji image analysis software. To verify HpaII digestion efficiency, pBluescript KS(+) plasmid was analysed by qPCR using one forward primer that is complementary to sequences upstream of the CCGG site of the β-lactamase gene (at 2580) and two different reverse primers that map upstream and downstream the HpaII sites. All analysed samples displayed 96–98% digestion efficiency. Bisulfite conversion was performed using the EpiTect Bisulfite Kit (Qiagen) according to the manufacturer's protocol. Amplification was performed with a nested touchdown PCR program with a 0.5 °C decrease in the annealing temperature after each cycle. Bisulfite primers are listed in Supplementary Table 7. Amplified sequences were cloned in pCR II-TOPO (Life Technologies) and sequenced. For COBRA assay, PCR products were digested with BstUI (New England Biolabs) followed by agarose gel electrophoresis. GlucMS-qPCR assay was performed using the EpiMark 5-hmC and 5mC-Analysis Kit (NEB). Briefly, genomic DNA was treated with T4 Phage β-glucosyltransferase. Glucosylated DNA was digested with HpaII or MspI or no enzyme (mock digestion) at 37 °C for 2 h and inactivated for 20 min at 80 °C. The HpaII- and MspI-resistant fraction was quantified by qPCR using primers designed around a single HpaII/MspI site, normalizing to the mock digestion control (Supplementary Table 7).

For promoter methylation analyses of Pramel7-regulated genes and ESCs versus 2i (ref. 12), WGBS data of E14 serum ESCs from ref. 8 were used. A minimum CpG coverage of 10 was set as a threshold. Average promoter methylation (+1 kb and −1 kb relative to the transcription start site) was calculated using deepTools[46] for all RefSeq RNA genes—retrieved through BioMart—taking each CpG into consideration (no binning). Random promoters from all RefSeq RNA genes were taken as a reference for average genome-wide promoter methylation using the same number of promoter regions as for upregulated gene sets together. Median promoter methylations with interquartile ranges are plotted for the indicated gene sets.

Yeast two-hybrid assay. Pramel7 was used as bait and a cDNA library of E14 ESCs as a pool of prey proteins. The bait was cloned into a lexA-expression vector and tested for self-activation and successful expression. cDNA library was transformed and co-expressed with the bait plasmid. Positive clones were selected and library plasmids were isolated. Finally positive clones were sequenced and BLAST analysis was performed.

Co-immunoprecipitation and proteomic analysis. Cell pellets were resuspended in IP buffer (50 mM Tris-HCl pH 7.5, 150 mM KCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 20% (v/v) glycerol, 0.5 mM dithiothreitol, 0.1% (v/v) NP-40, Proteinase inhibitor cocktail (Roche)), sonicated and DNase treated. One milligram of nuclear protein was subjected to immunoprecipitation overnight at 4 °C using ANTI-FLAG M2 affinity gel (Sigma). Precipitates were washed three times with IP buffer, separated on a 6% SDS-polyacrylamide gel and analysed by immunoblot. For proteomic analysis, immunoprecipitated proteins were eluted twice using 100 µl M2 FLAG-peptide (Sigma). Eluates were precipitated with trichloro-acetic acid (TCA), washed 4–6× with cold acetone and then dried at 95 °C. Proteomic analysis was carried out at the Functional Genomics Centre Zurich (University and ETH Zurich).

RNA extraction, reverse transcription and quantitative real-time PCR and RNA-seq. RNA was extracted using the RNeasy Mini Kit (Qiagen). Five hundred nanograms or 1 microgram of RNA was reverse transcribed using Oligo(dT)12–18 Primer (Life Technologies), 10 mM dNTP Mix (Life Technologies), RNAsin Plus RNAse Inhibitor (Promega) and SuperscriptIII Reverse Transcriptase (Life Technologies). Real-Time PCR was performed with the Rotor-Gene SYBR Green PCR Kit (FAST) (Qiagen). Primer sequences are listed in Supplementary Table 7.

Illumina RNA-sequencing. Data were deposited to the European Nucleotide Archive (ENA) under the accession number: PRJEB12665.

Library preparation. The quality of the isolated RNA was determined with a Qubit (1.0) Fluorometer (Life Technologies) and a Bioanalyzer 2100 (Agilent). Only those samples with a 260 nm/280 nm ratio between 1.8–2.1 and a 28S/18S ratio within 1.5–2 were further processed. The TruSeq RNA Sample Prep Kit v2 (Illumina) was used in the succeeding steps. Briefly, total RNA samples (100–1,000 ng) were ribo-depleted using Ribo Zero Gold(Epicentre) and then fragmented. The fragmented samples were reversed transcribed to cDNA, end-repaired and polyadenylated before ligation of TruSeq adapters containing the index for multiplexing. Fragments containing TruSeq adapters on both ends were selectively enriched with PCR. The quality and quantity of the enriched libraries were validated using Qubit (1.0) Fluorometer and the Caliper GX LabChip GX (Caliper Life Sciences). The product is a smear with an average fragment size of approximately 260 bp. The libraries were normalized to 10 nM in Tris-Cl 10 mM, pH 8.5 with 0.1% Tween-20.

Cluster generation and sequencing. The TruSeq PE Cluster Kit v4-cBot-HS or TruSeq SR Cluster Kit v4-cBot-HS (Illumina) was used for cluster generation using 10 pM of pooled normalized libraries on the cBOT. Sequencing was performed on the Illumina HiSeq 2500 paired end at 2 X101 bp or single end 100 bp using the TruSeq SBS Kit v4-HS (Illumina).

Data analysis. The raw reads were first cleaned by removing adapter sequences, trimming low-quality ends, and filtering reads with low quality (phred quality <20) using Trimmomatic[47]. Sequence alignment of the resulting high-quality reads to the *Mus musculus* reference genome (build GRCm38) was performed with STAR (version 2.4.2a) (ref. 48). The aligned reads were counted using the featureCount method of the R-package Rsubread (version 1.18.0). To detect differentially expressed genes we applied a count-based negative binomial model implemented in the software package edgeR (version: 3.10.2) (ref. 49), in which the normalization factor was calculated by the trimmed mean of M values (TMM) method[49]. The gene-wise dispersions were estimated by conditional maximum likelihood and an empirical Bayes procedure was used to shrink the dispersions towards a consensus value. The differential expression was assessed using an exact test adapted for over-dispersed data. Genes showing altered expression with adjusted (Benjamini and Hochberg method) P value < 0.05 were considered as differentially expressed.

Principal component analysis (PCA). The principal component analysis has been performed on the qRT–PCR data set combined with the RNA-seq data set described in ref. 9. The $\log_2$ counts in RNA-seq have been selected and normalized to mean. The data set taken for PCA and hierarchical clustering was created as a projection of

DOI: 10.1038/ncb3554

96 genes used in ref. 9 onto the Pramel7 ESCs RNA-seq experiment. The data sets of Borowiak *et al.* (ref. 9) and our own RNA-seq were combined and batch correction was applied using the ComBat function from the sva Bioconductor package[50]. PCA has been calculated using the FactoMineR[51] R library.

Transfection of HEK293T cells and electroporation of ESCs. Transfection of HEK293T with plasmid DNA was performed using either $CaCl_2$ and BES solution (50 mM BES, 280 mM NaCl, 1.5 mM Na_2HPO_4, pH 7.0) or Xtremegene HP (Roche) transfection reagent. For stable integration of the plasmid pCAG-FLAG-Pramel7-PGK-puro in E14 ESCs, the cells were first separated from the feeder layer by trypsinization, incubated with plasmid DNA and electroporated at 500 µF capacitance and 240 V using a Bio Rad Gene Pulser II. Antibiotic selection was started 24–48 h after electroporation and continued for 4 subsequent days (puromycin, 1 µg ml^{-1}). After the selection medium was changed to standard conditions (serum +LIF), the selected colonies were grown and clonally expanded.

For the time course experiment in HEK293T cells, HEK293T cells were transfected with FLAG-Pramel7 expression plasmid. Samples were collected 8, 12, 24, 36, 48, 72, 96 and 120 h after transfection and analysed for FLAG-Pramel7 and UHRF1 protein levels by western blotting. Endogenous ubiquitylation on UHRF1 was determined by anti-UHRF1 immunoprecipitation of HEK293 cells 36 h after transfection with UHRF1 and FLAG-Pramel7 expression plasmid. When indicated, plasmids expressing WT, Lys0 or Lys48 mutant FLAG-ubiquitin were co-transfected and ubiquitin was detected with FLAG antibodies. Proteasomal degradation of endogenous UHRF1 was monitored in 28 h post-transfection HEK293 cells treated with MG132 (20 µM) for 8 h.

Differentiation and reconversion of control E14 ESCs and P7 ESCs. E14 ESCs and P7 ESCs were separated from the feeder layer and seeded on gelatinized feeder-free culture dishes in complete medium without LIF. To avoid confluence, cells were passaged after 6 and 10 days. After 14 days, the cells were seeded on two 96-well plates containing feeder cells and LIF in a dilution of 10 cells per well and cultured for a further 7 days. Fourteen-day differentiated cells and cells cultured in the presence of LIF were harvested and analysed for the expression of Pramel7 and UHRF1 and of the pluripotency factors Oct4, Nanog, Rex1 and SSEA1 (immunofluorescence and qPCR).

CRISPR/CAS9 targeting of Pramel7 in preimplantation embryo. The Cas9 target sequence GCCTAAGAAGCAAATAGTGG in the Pramel7 gene was selected using the CRISPOR search algorithm http://crispor.tefor.net. The target sequence was incorporated into a dsDNA GenBlock fragment (IDT) 5′-GCGCGCTAATACGACTCACTATAGCCTAAGAAGCAAATAGTGGGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTT-3′ containing a 5′ proximal T7 promoter and a 3′ proximal sgRNA sequence shown to promote optimal assembly of Cas9 RNPs[52]. The sgRNA RNA was synthesized using the T7 transcription kit (NEB) following the manufacturer's instructions and the resulting transcripts were purified using the NuclAway gel filtration spin columns (Ambion). The Cas9 RNPs were assembled by heating the solution of 300 ng µl^{-1} sgRNA at 70 °C for 5 min, cooling to room temperature and incubating with 67 ng µl^{-1} of Cas9 protein (Toolgene) at room temperature for 20 min in a final volume of 30 µl. For the control Rosa26 RNP the procedure was repeated using a similar sgRNA targeting the XbaI site of the Rosa26 locus[53]. The Cas9 RNPs targeting the Pramel7 and Rosa26 were microinjected into the pronuclei of C57Bl/6J zygotes essentially as described in ref. 53.

We performed two independent injection rounds. The first round was performed with 224 embryos (Rosa26 = 114 embryos, P7 = 110 embryos), the second with 178 embryos (Rosa26 = 87 embryos, P7 = 91 embryos). Based on our experience, this number of embryos is sufficient to assess an effect. The surviving zygotes were cultured for 4 days in KSOM medium (Millipore). The development of the embryos was checked daily using an inverted microscope equipped with Nomarski optics. No statistical method was used to predetermine sample size and the experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Antibodies. For immunofluorescence analysis, the following antibodies were used: anti-5mC (1:1,000: Diagenode C15200081), anti-OCT-3/4 (1:500: Santa Cruz sc-9081), anti-SSEA1 (1:200: Developmental Studies Hybridoma Bank, University of Iowa), anti-UHRF1 (1:200: Santa Cruz sc-98817).

For western blotting and immunoprecipitation analyses the following antibodies were used: anti-H3 (1:15,000: Abcam ab1791), anti-Tubulin (1:10,000: Sigma T8203), anti-FLAG (1:15,000: Sigma F7425), anti-UHRF1 (1:3,000: Santa Cruz sc-98817), anti-Elongin B (1:1,000: Santa Cruz sc-11447), anti-Elongin C (1:1,000: Santa Cruz sc-1559), anti-Cullin 2 (1:2,000: Invitrogen 51-1800), anti-ubiquitin (1:1,000: Santa Cruz sc-8017). Pramel7 antibodies were produced in rabbit through immunization with a specific Pramel7 peptide (mPRAMEL7-5: NH2-CRDYLVGTLPKKQIVEDHSR-COOH, Pineda Antikörper Services). Serum was purified by Protein A Sepharose followed by affinity chromatography and used with a dilution of 1:5,000 for western blotting and 1:200 for embryo staining.

Statistics and reproducibility. For animal experiments no statistical method was used to predetermine sample size and the experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment. RNA-Seq analysis was performed in triplicates: genes showing altered expression with adjusted (Benjamini and Hochberg method) P value < 0.05 were considered as differentially expressed. Bar graphs represent mean ± s.d.

Data availability. RNA-seq data that support the findings in Fig. 1e,f and Supplementary Fig. 2 and Supplementary Tables 1–4 have been deposited in the European Nucleotide Archive (ENA) under the accession code PRJEB12665. Previously published RNA-seq and Bisulfite-seq data that were reanalysed here are available under accession codes GSE20187, GSE23943 and GSE41923 at the Gene Expression Omnibus (Figs 1e and 2g and Supplementary Tables 2–4), in the ArrayExpress repository under accession number E-MTAB-2555 (Fig. 1f and Supplementary Fig. 2) and in the European Nucleotide Archive as study ERP005749 (Fig. 1f and Supplementary Fig. 2). Source data for Figs 1b, 4c, 6 and 8a; and Supplementary Figs 1a, 4b and 5b have been provided as Supplementary Table 8. All other data supporting the findings of this study are available from the corresponding authors on reasonable request.

45. Tsumura, A. *et al.* Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805–814 (2006).
46. Ramírez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
47. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
48. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
49. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
50. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
51. Le, S., Josse, J. & Husson, F. FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18 (2008).
52. Chen, B. *et al.* Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479–1491 (2013).
53. Hermann, M., Cermak, T., Voytas, D. F. & Pelczar, P. Mouse genome engineering using designer nucleases. *J. Vis. Exp.* **86**, 50930 (2014).

Corrigendum: Pramel7 mediates ground-state pluripotency through proteasomal–epigenetic combined pathways

Urs Graf, Elisa A. Casanova, Sarah Wyck, Damian Dalcher, Marco Gatti, Eva Vollenweider, Michal J. Okoniewski, Fabienne A. Weber, Sameera S. Patel, Marc W. Schmid, Jiwen Li, Jafar Sharif, Guido A. Wanner, Haruhiko Koseki, Jiemin Wong, Pawel Pelczar, Lorenza Penengo, Raffaella Santoro and Paolo Cinelli

Nature Cell Biology http://dx.doi.org/10.1038/ncb3554 (2017); published online 12 June 2017; corrected after print 21 June 2017

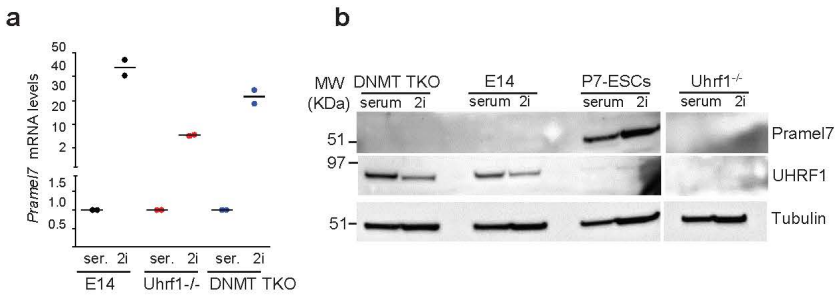
In the version of this Article originally published, the following affiliation was omitted for Sarah Wyck: Clinic of Reproductive Medicine, University of Zurich, Winterthurerstrasse 260, CH-8057 Zurich, Switzerland. This has been corrected in the online version of the Article.

SUPPLEMENTARY INFORMATION

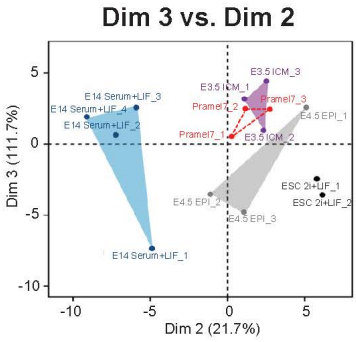
nature cell biology

DOI: 10.1038/ncb3554

In the format provided by the authors and unedited.

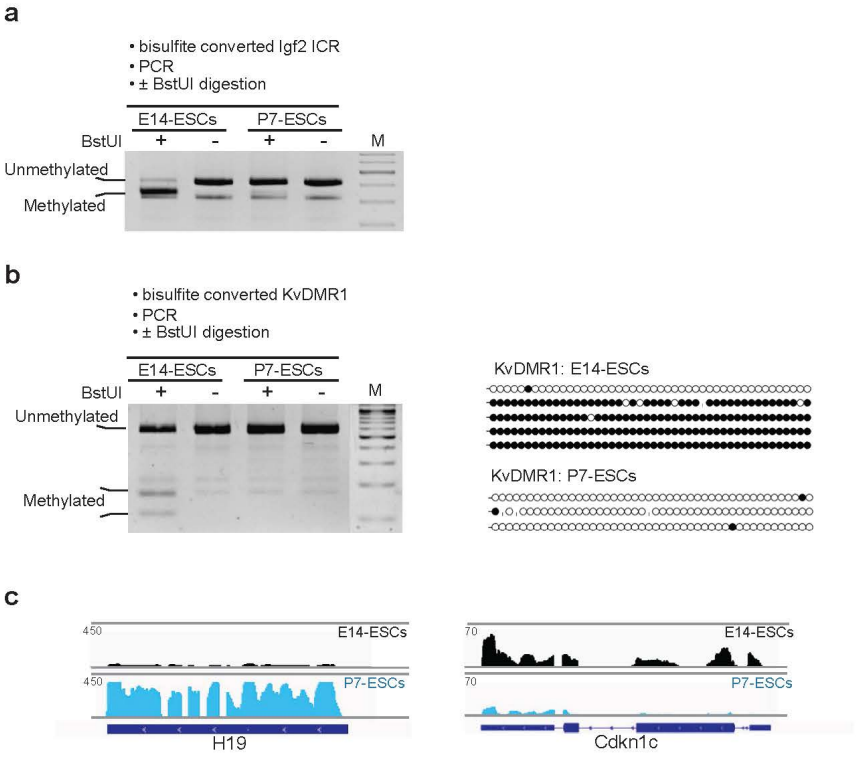


Supplementary Figure 1 (Related to Fig. 1) **a**, Expression of Pramel7 in E14, *Uhrf1*$^{-/-}$ and DNMT-TKO ESCs grown in serum/LIF or 2i conditions. Pramel7 mRNA levels were measured by qRT-PCR and normalized to Actin mRNA. Average values of two independent experiments, lines represent means. **b**, Western blot showing expression levels of Pramel7 and UHRF1 in E14, DNMT-TKO, *Uhrf1*$^{-/-}$ ESCs and P7-ESCs cultured in serum/LIF and 2i conditions detected with Pramel7 and UHRF1 antibodies. Tubulin is shown as loading control. Representative of 3 independent experiments. Unprocessed original scans of immunoblots are shown in Supplementary Figure 6.



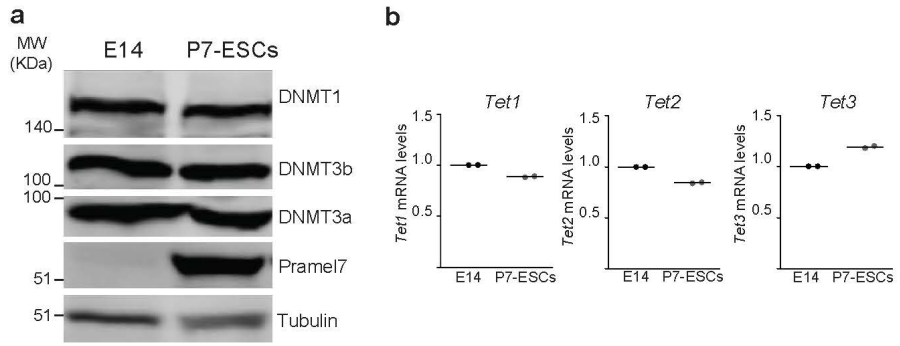
Supplementary Figure 2 (Related to Fig. 1) Correlation of P7-ESC gene expression to the early embryo stages. PCA analysis (dimensions 2 and 3) of P7-ESCs, embryonic stages E3.5 and E4.5, ESCs cultured in serum/LIF or 2i from data set published in [9].

SUPPLEMENTARY INFORMATION



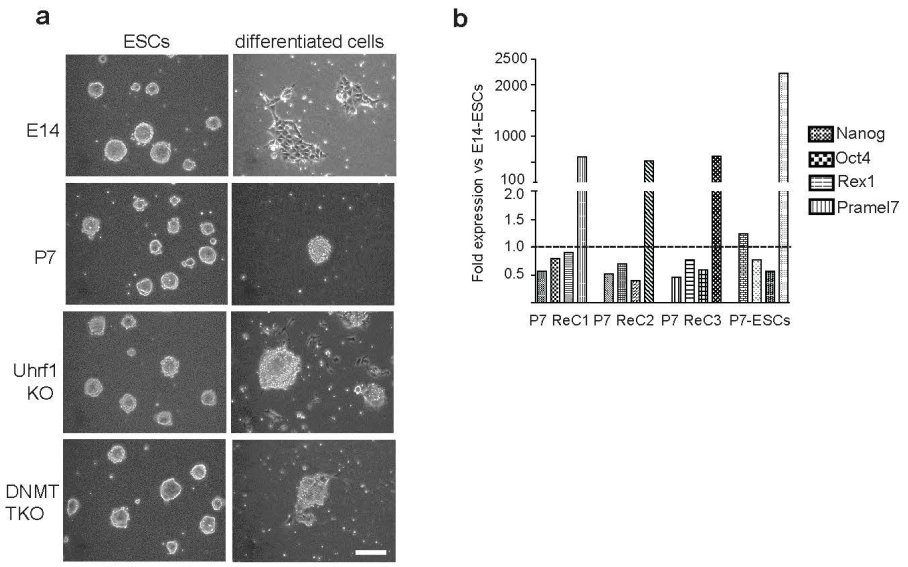
Supplementary Figure 3 (Related to Fig. 2) **a, b,** COBRA analysis showing extensive demethylation of *H19/Igf2* ICR and KvDMR1 (an intronic CpG island within the *KCNQ1* gene) in P7-ESCs. Bisulfite converted DNA was amplified with bisulfite specific primers and methylation content was assessed by digestion with BstUI that recognizes CGCG sequences. Digestion of BstUI serves as indicator of the presence of methylated sequences which were resistant to C to U conversion. Bisulfite sequencing of KvDMR1 (**b**) supports the lack of DNA methylation measured by COBRA assay. **c,** Example of changes in transcription of two imprinting regulated genes in P7-ESCs. Experiments in **a** and **b** were repeated independently twice.

SUPPLEMENTARY INFORMATION



Supplementary Figure 4 (Related to Fig. 2) **a,** Western blot showing expression levels of DNMT1, 3a and 3b in E14 and P7-ESCs cultured in serum/LIF. Tubulin is shown as loading control. **b,** *Tet1*, *Tet2* and *Tet3* expression levels measured by RT-qPCR. Values were normalized to Actin mRNA and represent the average of two independent experiments. Average values of two independent experiments. Unprocessed original scans of immunoblots are shown in Supplementary Figure 6.

SUPPLEMENTARY INFORMATION



Supplementary Figure 5 (Related to Fig. 7 and 8) **a**, Representative images of ESCs grown in 2i condition and upon 5 days of differentiation. Upon differentiation methylation defective *Uhrf1*$^{-/-}$ and DNMT-TKO ESCs formed less differentiated and more compact colonies as in the case of P7-ESCs. Scale bar: 200µm. Representative of 2 independent experiments. **b**, Expression of pluripotency-associated genes *Nanog*, *Oct4* and *Rex1* in three P7 reconverted clones. Values were measured by qRT-PCR, normalized to *actin* mRNA and represented as fold change relative to the expression in E14-ESCs. Average values of two independent experiments.

4. Discussion

Ground state ESCs cultured with LIF and 2i have been suggested a good tool for studying ground state pluripotency, since ESCs+2i represent a homogenous population highly resembling the in vivo pre-implantation epiblast cells of the developing embryo (Boroviak et al., 2014; Hackett and Surani, 2014). Here we highlight fundamental differences in the chromatin organization and its dependency on several chromatin regulators in ground state and primed ESCs. We focused particularly on the nucleolar remodeling complex (NoRC) and showed that only ground state ESCs are dependent on TIP5, the main component of NoRC. TIP5 is highly expressed in ESCs but not bound to rRNA genes, as found in differentiated cells. Upon TIP5 depletion, ground state ESCs grow slower and are highly impaired in their differentiation potential. Furthermore, we demonstrated that NoRC is required for maintaining correct gene expression signatures and epigenetic landscapes in ESCs+2i. Importantly, this strong dependency on TIP5 was ground state-specific since TIP5 depletion in ESCs+serum did not result in any detectable defects so far. Strikingly, we could recapitulate highly similar, ground state-specific phenotypes by depleting or inactivating other chromatin regulators such as TOP2A and cohesin, which associate with TIP5 on chromatin. HiC analyses revealed an essential role of TIP5 in restricting far-cis chromatin contacts in ground state ESCs. Taken together, these findings imply that chromatin structures must be held under tight control by several chromatin regulators for maintaining ground state pluripotency. ESCs+serum by contrast, are developmentally primed and seem to be less dependent on chromatin regulators for reinforcing active and repressed genome partitioning.

4.1 Ground state ESC chromatin structures

Here we described a ground state-specific function of NoRC in regulating the gene expression signatures and epigenetic landscapes. TIP5 expression and its global genomic binding pattern were very similar in ESCs+2i and ESCs+serum, implying that it is most likely not NoRC itself but rather the state of its substrate that determines its functional relevance in ESCs. Global genome hypomethylation and low H3K27me3 levels at PRC-target sites have been among the most investigated chromatin characteristics identified for ESCs+2i (Ficz et al., 2013; Habibi et al., 2013; Marks et al., 2012). However, analyses of TIP5 depletion in PRC component (Eed

and Ring1b) KOs and DNA methylation triple KOs (TKOs) did not reveal any evidences that NoRC function might be linked via these two repressive pathways. Recently, ultra-high resolution microscopy by STORM investigating the chromatin structures at the nucleosome level revealed that ESCs+2i harbor a more open chromatin configuration when compared to ESCs+serum. So-called "nucleosome clutches" were significantly smaller in size and contained less nucleosomes per clutch in ESCs+2i (Ricci et al., 2015). Here we propose a model in which the open and dynamic nature of ground state chromatin might predispose ESCs+2i to a dependency on chromatin regulators. ESC+serum, containing a more restricted chromatin state including pre-established long-range interactions (Joshi et al., 2015), might be less dependent on the contribution of mechanisms monitoring chromatin structures. In line with this model, we showed that depletion or chemical inactivation of two additional chromatin structure regulators, namely TOP2A and cohesin, could highly recapitulate the alterations observed upon TIP5 depletion in ESCs+2i but not in ESCs+serum. Alternatively, we cannot exclude that posttranslational modifications restricting NoRC activity would be responsible for the functional differences observed in ground state and primed ESCs. In fact, a recent study indeed identified S1377 on TIP5 to be differentially phosphorylated between ESCs+2i and ESCs+serum (Rafiee et al., 2016). Whether this phosphorylation site may influence NoRC activity remains to be tested experimentally. However, differences in NoRC activity mediated by PTMs alone could not explain that chemical inhibition or depletion of TOP2A highly recapitulates the ground state-specific defects that were observed in TIP5 depleted ESCs+2i.

4.2 Altered chromatin structures caused by TIP5 depletion independently affect gene expression and H3K27me3 occupancies in ESCs+2i

The fact that TIP5 binds to large and active compartments of the genome (type A) and that this occupancy does not correlate with sites of TIP5-regulated genes led us hypothesize that NoRC might be rather involved in regulating 3D structures in ground state ESCs. Indeed, we identified several chromatin regulators directly interacting with TIP5 on chromatin in ESCs. Furthermore, HiC analyses revealed a significant increase in far-cis contacts upon TIP5 depletion in ESCs+2i. These observations suggest that the changes in gene expression and H3K27me3 are both consequences of

altered chromatin structures caused by a loss in NoRC function in ESCs+2i. We would like to further emphasize that the changes in gene expression and H3K27me3 do not seem to be directly coupled. Although differently expressed genes were enriched in H3K27me3, TIP5 depletion in Eed and Ring1b KO in ESCs+2i revealed highly similar gene expression changes as observed in wt ESCs+2i. Thus, PRC-mediated repression is not implicated in TIP5-mediated regulation of gene expression. These results are also in agreement with a recent study demonstrating that global expression of bivalent genes is not altered in the absence of H3K27me3 (Galonska et al., 2015b). Thus, it rather seems that the gene expression changes and global H3K27me3 redistribution display two independent read outs of altered chromatin structures triggered by TIP5 depletion. Moreover, we observed that the genes affected by TIP5 knockdown are also often differentially expressed in ESCs+2i and ESCs+serum, suggesting that these genes – so far for unknown reasons – are in general more sensitive to changes in chromatin structures in ESCs.

4.3 3D chromatin architecture and gene regulation

3D compartmentalization of the genome within the nucleus has been proposed to be highly important for gene regulatory circuits (Bonev et al., 2017). CTCF and cohesin shape the genome organization into so called topological associated domains (TADs) (Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017). While CTCF recognizes specific DNA sequence motif and functions as an insulating boundary, cohesin is thought to be loaded at sites of active transcription and travel along chromosome arms by a loop extrusion mechanism until reaching convergently oriented CTCF sites (de Wit et al., 2015; Guo et al., 2015; Nora et al., 2017; Sanborn et al., 2015). In comparison to ESCs+serum, ESCs+2i contain more CTCF occupied sites and loop structures (Jonathan A. Beagan, 2017), further suggesting that the 3D chromatin organization might play a particularly important role in ground state ESCs for maintaining gene regulatory mechanisms. Indeed, our experiments highlighted that cohesin depletion – known to cause a loss in loop structures (Rao et al., 2017) – leads to more pronounced gene expression changes in ESCs+2i when compared to ESCs+serum.

Our proteomic interaction studies revealed that the highest interacting partner of TIP5 is SNF2h, suggesting that the function of TIP5 in ESCs is directed via the NoRC

complex. Furthermore, this analysis showed that TIP5 tightly associates with TOP2A and cohesin on chromatin. Strikingly, we did not detect CTCF as TIP5-interacting partner in any of our mass spectrometric analyses. This implies that cohesin associates with NoRC while traveling along the chromosome arms rather than at the stable anchor point that are co-occupied by CTCF. Accordingly, we did not detect TIP5 enrichment at CTCF/cohesin peaks. Our HiC analysis revealed global changes in far-cis contacts (>10Mb) but does not provide the required resolution for elucidating changes in local chromatin structures. However, a recent study highlighted that abolishing cohesin loading resulted in a highly similar increase in these long-range chromatin contacts (Haarhuis et al., 2017). Since cohesin is known to limit local chromatin dynamics (Nozaki et al., 2017), these findings suggested that local chromatin compaction might globally limit far-cis chromatin contacts and therefore reinforce active and repressed genome partitioning (Haarhuis et al., 2017; Schwarzer et al., 2017). Strikingly, the reciprocal experiment of restricting cohesin removal from chromatin revealed the inverse outcome resulting in a reduction of far-cis chromatin contacts, further supporting this hypothesis (Haarhuis et al., 2017). Taking these assumptions into consideration for a putative role of TIP5 in ESCs leads to the intriguing model by which NoRC might mediate local chromatin compaction and thereby limit far-cis contacts in ground state ESCs. Chromatin of ESCs+serum by contrast, generally contains more condensed chromatin structures at the level of nucleosomes (Ricci et al., 2015). Thus, primed ESCs might be less dependent on factors that additionally limit chromatin dynamics and flexibility, explaining why ESCs+serum are less dependent on NoRC than ground state ESCs+2i. How NoRC mechanistically controls local chromatin dynamics remains elusive. Our HiC experiments did not indicate any global changes in TAD organization, arguing against a putative role of NoRC in regulating global cohesin occupancies. Though, it remains to be determined whether NoRC might control individual loop structures at a subTAD organization level. In fact, one major question in the field of chromatin architecture aims to elucidate what controls stable loop formation between two CTCF-bound sites. A vast majority of chromosomal contacts between CTCF-bound sites occur at a short range (<1Mb) (Handoko et al., 2011; Rao et al., 2014; Tang et al., 2015). Yet, only a small subset of all CTCF-bound sites mediate chromosomal contacts, implying that cells must have evolved mechanisms that control on and off rates between the interaction of CTCF and the extruding cohesin complex (Barbieri et al., 2017; Rao et

al., 2014). Interestingly, visual inspection of the HiC contact maps revealed prominent changes at many subTAD loop structures in TIP5-depleted ESCs+2i. The spatial organization of the HoxA gene cluster for instance, underwent an entire re-organization in ESCs+2i lacking TIP5. A subTAD demarcating this locus completely dissolved and fused with two neighboring subTAD structures creating a new subTAD boundary. Strikingly, genes located upstream of this new boundary showed transcriptional upregulation, whereas downstream genes were tendentially downregulated in response to TIP5 knockdown. Taken together, these observations highlight a putative role of NoRC in regulating stable loop associations. In line with this hypothesis, SNF2H – the catalytic subunit of NoRC – has been shown to specifically regulate nucleosome occupancies at CTCF sites in human cells (Wiechens et al., 2016). Future high-resolution chromatin contact maps determined by 5C experiments might shed further light on the exact mechanistic role of NoRC at these loci.

4.4 Topoisomerases and supercoiling in genome architecture

While cohesin and its contribution to 3D genome organization into TADs has been extensively studied in the past years (Rao et al., 2017; Schwarzer et al., 2017), the role of topoisomerases in regulating chromatin structures and genome architecture in mammalian cells remains under-investigated. Studies in ESCs+serum suggested that TOP2A chromatin binding is partially directed via the BAF nucleosome-remodeling complex (Dykhuizen et al., 2013). Here, we additionally identified NoRC as an interactor of TOP2A in mouse ESCs. Our results suggest that they synergize in order to maintain correct gene expression patterns and epigenetic signatures in ground state ESCs. Strikingly, TOP2A genomic binding occurs at active sites of the genome that poorly correlate with H3K27me3 (Dykhuizen et al., 2013; Thakurela et al., 2013), a binding profile that is similar to TIP5 genome occupancy defined by our analyses. Unfortunately, due to the discontinuation of validated high-grade ChIP TOP2a antibodies (Naughton et al., 2013; Thakurela et al., 2013), we have been unable to measure TOP2A genome occupancy in ESCs+2i.

Computational modeling was used to propose an active role of topoisomerases in shaping 3D chromatin structures. In their simulations, loop structures alone could not account for contact frequencies that had been experimentally observed within TADs.

Only by adding supercoiling as an additional variable, they could recapitulate the degree of insulation among different TADs. Specifically long distance enhancer-promoter contacts were predicted to be stabilized by supercoiling (Benedetti et al., 2014a; Benedetti et al., 2014b). Furthermore, Stasiak and colleagues have recently proposed an additional link between topoisomerases and genome organization (Racko et al., 2017). In their mechanistic model, transcription-induced supercoiling catalyzed by TOP1 could provide the force for pushing handcuff-shaped cohesin rings along chromosome arms. It was further proposed that TOP2B – that has been shown to bind CTCF-occupied sites at TAD borders in somatic cells (Uuskula-Reimand et al., 2016) – could resolve the torsional stress leading to the stable association of cohesin at TAD boundaries (Racko et al., 2017). In summary, topoisomerases have been implicated in several processes regulating genome organization in mammalian cells. Thus, it is likely that its depletion causes severe effects in chromatin structures. The striking similarity in the phenotypes observed in TIP5 and TOP2A depletion experiments indicated that both factors are likely involved in similar pathways. This hypothesis is further supported by the fact that their co-depletion did not reveal any additive effects in gene expression changes. Importantly, knockdown of these two factors did not result in any detectable defects in gene expression or H3K27me3 in ESCs+serum, further strengthening the idea that ground state ESCs are particularly dependent on this set of chromatin regulators. In summary, our data suggests that ground state ESC chromatin – more open and dynamic in its nature – strictly requires mechanisms that limit local chromatin contacts to maintain proper gene expression patterns and epigenetic landscapes.

4.5 Switch of NoRC function according to developmental and pathological state

Previous results from our laboratory determined that the functional role of NoRC highly depends on the developmental and pathological state of cells. In somatic and healthy cells TIP5 is exclusively localized within the nucleolus, the compartment where rRNA genes reside and represses their transcription (Guetg et al., 2012; Santoro et al., 2002). By contrast, in both ground state and primed ESCs TIP5 is highly expressed but does not associate and hence, does not regulate transcription of rRNA genes (this work and (Leone et al., 2017; Savic et al., 2014). Upon early onset

of differentiation, NoRC is rapidly recruited to a subset of rRNA genes and progressively mediates their epigenetic silencing (Leone et al., 2017; Savic et al., 2014). Here we describe a further switch in NoRC function from ground state to primed pluripotent state, a system that has been used to model very early steps of embryonic development (Atlasi and Stunnenberg, 2017; Joshi et al., 2015). While NoRC controls expression of many developmental genes and regulates genome-wide H3K27me3 occupancy in ground state pluripotency, it remains elusive whether NoRC has any functional relevance in primed ESCs. However, the switch to regulating rRNA genes occurs only upon exit of pluripotency, since NoRC neither binds nor silences rDNA in ESCs+serum yet. Interestingly, NoRC also changes its function during the transition from healthy to cancer cells (Gu et al., 2015a). Similarly to ESCs, TIP5 is abundantly present in the nucleoplasm in metastatic prostate cancer (PCa) cells and regulates genes involved in developmental processes that are frequently repressed in metastatic PCa (Gu et al., 2015b). Considering that cancer cells often acquire stem cell-like features, it is likely that TIP5 might use similar mechanisms to regulate ESCs and cancer cells. Taken together, we demonstrate that highly related cell types (ground state and primed ESCs) markedly differ in their dependency on various chromatin regulators. This suggests that the global shape of chromatin is a fundamental player for gene regulation and chromatin states and highlights different degrees of genome plasticity during the early phases of development.

5. References

Abranches, E., Bekman, E., and Henrique, D. (2013). Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression. PloS one *8*, e59928.

Ahmed, K., Dehghani, H., Rugg-Gunn, P., Fussner, E., Rossant, J., and Bazett-Jones, D.P. (2010). Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. PloS one *5*, e10531.

Akhtar, W., de Jong, J., Pindyurin, A.V., Pagie, L., Meuleman, W., de Ridder, J., Berns, A., Wessels, L.F., van Lohuizen, M., and van Steensel, B. (2013). Chromatin position effects assayed by thousands of reporters integrated in parallel. Cell *154*, 914-927.

Akimitsu, N., Adachi, N., Hirai, H., Hossain, M.S., Hamamoto, H., Kobayashi, M., Aratani, Y., Koyama, H., and Sekimizu, K. (2003). Enforced cytokinesis without complete nuclear division in embryonic cells depleting the activity of DNA topoisomerase IIalpha. Genes Cells *8*, 393-402.

Alexandrova, S., Kalkan, T., Humphreys, P., Riddell, A., Scognamiglio, R., Trumpp, A., and Nichols, J. (2016). Selection and dynamics of embryonic stem cell integration into early mouse embryos. Development *143*, 24-34.

Aloia, L., Di Stefano, B., and Di Croce, L. (2013). Polycomb complexes in stem cells and embryonic development. Development *140*, 2525-2534.

Alvarez-Saavedra, M., De Repentigny, Y., Lagali, P.S., Raghu Ram, E.V., Yan, K., Hashem, E., Ivanochko, D., Huh, M.S., Yang, D., Mears, A.J.*, et al.* (2014). Snf2h-mediated chromatin organization and histone H1 dynamics govern cerebellar morphogenesis and neural maturation. Nature communications *5*, 4181.

Amendola, M., and van Steensel, B. (2015). Nuclear lamins are not required for lamina-associated domain organization in mouse embryonic stem cells. EMBO reports *16*, 610-617.

Amouroux, R., Nashun, B., Shirane, K., Nakagawa, S., Hill, P.W., D'Souza, Z., Nakayama, M., Matsuda, M., Turp, A., Ndjetehe, E.*, et al.* (2016). De novo DNA methylation drives 5hmC accumulation in mouse zygotes. Nature cell biology *18*, 225-233.

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., and Shirakawa, M. (2008). Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature *455*, 818-821.

Atlasi, Y., and Stunnenberg, H.G. (2017). The interplay of epigenetic marks during stem cell differentiation and development. Nature reviews Genetics *18*, 643-658.

Austenaa, L.M., Barozzi, I., Simonatto, M., Masella, S., Della Chiara, G., Ghisletti, S., Curina, A., de Wit, E., Bouwman, B.A., de Pretis, S.*, et al.* (2015). Transcription of Mammalian cis-Regulatory Elements Is Restrained by Actively Enforced Early Termination. Molecular cell *60*, 460-474.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes & development *17*, 126-140.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M.*, et al.* (2006). Chromatin signatures of pluripotent cell lines. Nature cell biology *8*, 532-538.

Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Herault, Y., Guillou, F., and Bourc'his, D. (2016). The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science *354*, 909-912.

Barbieri, M., Xie, S.Q., Torlai Triglia, E., Chiariello, A.M., Bianco, S., de Santiago, I., Branco, M.R., Rueda, D., Nicodemi, M., and Pombo, A. (2017). Active and poised promoter states drive folding of the extended HoxB locus in mouse embryonic stem cells. Nat Struct Mol Biol *24*, 515-524.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schubeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature.

Beagan, J.A., Duong, M.T., Titus, K.R., Zhou, L., Cao, Z., Ma, J., Lachanski, C.V., Gillis, D.R., and Phillips-Cremins, J.E. (2017). YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. Genome Res *27*, 1139-1152.

Bell, P., Dabauvalle, M.C., and Scheer, U. (1992). In vitro assembly of prenucleolar bodies in Xenopus egg extract. J Cell Biol *118*, 1297-1304.

Benedetti, F., Dorier, J., Burnier, Y., and Stasiak, A. (2014a). Models that include supercoiling of topological domains reproduce several known features of interphase chromosomes. Nucleic acids research *42*, 2848-2855.

Benedetti, F., Dorier, J., and Stasiak, A. (2014b). Effects of supercoiling on enhancer-promoter contacts. Nucleic acids research *42*, 10425-10432.

Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R.*, et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bhattacharya, D., Talwar, S., Mazumder, A., and Shivashankar, G.V. (2009). Spatio-temporal plasticity in chromatin organization in mouse cell differentiation and during Drosophila embryogenesis. Biophysical journal *96*, 3832-3839.

Bian, Q., Khanna, N., Alvikas, J., and Belmont, A.S. (2013). beta-Globin cis-elements determine differential nuclear targeting through epigenetic modifications. J Cell Biol *203*, 767-783.

Bibel, M., Richter, J., Schrenk, K., Tucker, K.L., Staiger, V., Korte, M., Goetz, M., and Barde, Y.A. (2004). Differentiation of mouse embryonic stem cells into a defined neuronal lineage. Nat Neurosci *7*, 1003-1009.

Bochar, D.A., Savard, J., Wang, W., Lafleur, D.W., Moore, P., Cote, J., and Shiekhattar, R. (2000). A family of chromatin remodeling factors related to Williams syndrome transcription factor. Proceedings of the National Academy of Sciences of the United States of America *97*, 1038-1043.

Boettiger, A.N., Bintu, B., Moffitt, J.R., Wang, S., Beliveau, B.J., Fudenberg, G., Imakaev, M., Mirny, L.A., Wu, C.T., and Zhuang, X. (2016). Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. Nature *529*, 418-422.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. Nature reviews Genetics *17*, 661-678.

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.P., Tanay, A.*, et al.* (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell *171*, 557-572 e524.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nature cell biology *16*, 516-528.

Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science *317*, 1760-1764.

Boveri, T. (1909). Die blastomerenkerne von Ascaris meglocephala und die
theorie der chromosomenindividualitaet. Archiv für Zellforschung *3, 181–268.*

Bozhenok, L., Wade, P.A., and Varga-Weisz, P. (2002). WSTF-ISWI chromatin remodeling complex targets heterochromatic replication foci. The EMBO journal *21*, 2231-2241.

Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A.*, et al.* (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448*, 191-195.

Brookes, E., de Santiago, I., Hebenstreit, D., Morris, K.J., Carroll, T., Xie, S.Q., Stock, J.K., Heidemann, M., Eick, D., Nozaki, N.*, et al.* (2012). Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ESCs. Cell stem cell *10*, 157-170.

Buehr, M., Meek, S., Blair, K., Yang, J., Ure, J., Silva, J., McLay, R., Hall, J., Ying, Q.L., and Smith, A. (2008). Capture of authentic embryonic stem cells from rat blastocysts. Cell *135*, 1287-1298.

Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Randazzo, F., Metzger, D., Chambon, P., Crabtree, G.*, et al.* (2000). A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. Molecular cell *6*, 1287-1295.

Burton, A., and Torres-Padilla, M.E. (2014). Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. Nature reviews Molecular cell biology *15*, 723-734.

Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., and Peters, J.M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. Nature *544*, 503-507.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Molecular cell *49*, 825-837.

Carter, A.C., Davis-Dusenbery, B.N., Koszka, K., Ichida, J.K., and Eggan, K. (2014). Nanog-independent reprogramming to iPSCs with canonical factors. Stem Cell Reports *2*, 119-126.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230-1234.

Chambers, I., and Tomlinson, S.R. (2009). The transcriptional foundation of pluripotency. Development *136*, 2311-2322.

Chen, T., and Dent, S.Y. (2014). Chromatin modifiers and remodellers: regulators of cellular differentiation. Nature reviews Genetics *15*, 93-106.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol Cell Biol *23*, 5594-5605.

Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science *277*, 1996-2000.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu Rev Biochem *78*, 273-304.

Clapier, C.R., Iwasa, J., Cairns, B.R., and Peterson, C.L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nature reviews Molecular cell biology *18*, 407-422.

Cremer, T., and Cremer, M. (2010). Chromosome territories. Cold Spring Harbor perspectives in biology *2*, a003889.

de Dieuleveult, M., Yen, K., Hmitou, I., Depaux, A., Boussouar, F., Dargham, D.B., Jounier, S., Humbertclaude, H., Ribierre, F., Baulard, C.*, et al.* (2016). Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. Nature.

de Wit, E., Bouwman, B.A., Zhu, Y., Klous, P., Splinter, E., Verstegen, M.J., Krijger, P.H., Festuccia, N., Nora, E.P., Welling, M.*, et al.* (2013). The pluripotent genome in three dimensions is shaped around pluripotency factors. Nature *501*, 227-231.

de Wit, E., Vos, E.S., Holwerda, S.J., Valdes-Quezada, C., Verstegen, M.J., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H., and de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. Molecular cell *60*, 676-684.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes & development *25*, 1010-1022.

Dekker, J., and Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. Cell *164*, 1110-1121.

Dev, V.G., Tantravahi, R., Miller, D.A., and Miller, O.J. (1977). Nucleolus organizers in Mus musculus subspecies and in the RAG mouse cell line. Genetics *86*, 389-398.

Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. Molecular cell *62*, 668-680.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Doble, B.W., Patel, S., Wood, G.A., Kockeritz, L.K., and Woodgett, J.R. (2007). Functional redundancy of GSK-3alpha and GSK-3beta in Wnt/beta-catenin signaling shown by using an allelic series of embryonic stem cell lines. Developmental cell *12*, 957-971.

Domcke, S., Bardet, A.F., Adrian Ginno, P., Hartl, D., Burger, L., and Schubeler, D. (2015). Competition between DNA methylation and transcription factors determines binding of NRF1. Nature *528*, 575-579.

Du, Z., Song, J., Wang, Y., Zhao, Y., Guda, K., Yang, S., Kao, H.Y., Xu, Y., Willis, J., Markowitz, S.D*., et al.* (2010). DNMT1 stability is regulated by proteins coordinating deubiquitination and acetylation-driven ubiquitination. Science signaling *3*, ra80.

Du, Z., Zheng, H., Huang, B., Ma, R., Wu, J., Zhang, X., He, J., Xiang, Y., Wang, Q., Li, Y*., et al.* (2017). Allelic reprogramming of 3D chromatin architecture during early mammalian development. Nature *547*, 232-235.

Dundr, M., Misteli, T., and Olson, M.O. (2000). The dynamics of postmitotic reassembly of the nucleolus. J Cell Biol *150*, 433-446.

Dunn, S.J., Martello, G., Yordanov, B., Emmott, S., and Smith, A.G. (2014). Defining an essential transcription factor program for naive pluripotency. Science *344*, 1156-1160.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell systems *3*, 95-98.

Dykhuizen, E.C., Hargreaves, D.C., Miller, E.L., Cui, K., Korshunov, A., Kool, M., Pfister, S., Cho, Y.J., Zhao, K., and Crabtree, G.R. (2013). BAF complexes facilitate decatenation of DNA by topoisomerase IIalpha. Nature *497*, 624-627.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D., Buetow, K.H*., et al.* (2008). Global transcription in pluripotent embryonic stem cells. Cell stem cell *2*, 437-447.

Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., and Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic acids research *10*, 2709-2721.

Engel, W., Zenzes, M.T., and Schmid, M. (1977). Activation of mouse ribosomal RNA genes at the 2-cell stage. Human Genetics *38*, 57-63.

Erdel, F., and Rippe, K. (2011). Chromatin remodelling in mammalian cells by ISWI-type complexes--where, when and why? Febs j *278*, 3608-3618.

Esteve, P.O., Chin, H.G., Benner, J., Feehery, G.R., Samaranayake, M., Horwitz, G.A., Jacobsen, S.E., and Pradhan, S. (2009). Regulation of DNMT1 stability through SET7-mediated lysine methylation in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America *106*, 5076-5081.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Fatemi, M., and Wade, P.A. (2006). MBD family proteins: reading the epigenetic code. Journal of cell science *119*, 3033-3037.

Faunes, F., Hayward, P., Descalzo, S.M., Chatterjee, S.S., Balayo, T., Trott, J., Christoforou, A., Ferrer-Vaquer, A., Hadjantonakis, A.K., Dasgupta, R*., et al.* (2013). A membrane-associated beta-catenin/Oct4 complex correlates with ground-state pluripotency in mouse embryonic stem cells. Development *140*, 1171-1183.

Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. Development *121*, 273-285.

Fazzio, T.G., Huff, J.T., and Panning, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell *134*, 162-174.

Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. Nature *421*, 448-453.

Ferrai, C., Torlai Triglia, E., Risner-Janiczek, J.R., Rito, T., Rackham, O.J., de Santiago, I., Kukalev, A., Nicodemi, M., Akalin, A., Li, M.*, et al.* (2017). RNA polymerase II primes Polycomb-repressed developmental genes throughout terminal neuronal differentiation. Mol Syst Biol *13*, 946.

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., and Chambers, I. (2012). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. Cell stem cell *11*, 477-490.

Ficz, G., Hore, T.A., Santos, F., Lee, H.J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y.L., Walter, J.*, et al.* (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. Cell stem cell *13*, 351-359.

Finlan, L.E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J.R., and Bickmore, W.A. (2008). Recruitment to the nuclear periphery can alter expression of genes in human cells. PLoS Genet *4*, e1000039.

Fouse, S.D., Shen, Y., Pellegrini, M., Cole, S., Meissner, A., Van Neste, L., Jaenisch, R., and Fan, G. (2008). Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell stem cell *2*, 160-169.

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. The Journal of biological chemistry *278*, 4035-4040.

Gal-Yam, E.N., Jeong, S., Tanay, A., Egger, G., Lee, A.S., and Jones, P.A. (2006). Constitutive nucleosome depletion and ordered factor assembly at the GRP78 promoter revealed by single molecule footprinting. PLoS Genet *2*, e160.

Galonska, C., Ziller, M.J., Karnik, R., and Meissner, A. (2015a). Ground State Conditions Induce Rapid Reorganization of Core Pluripotency Factor Binding before Global Epigenetic Reprogramming. Cell Stem Cell *17*, 462-470.

Galonska, C., Ziller, M.J., Karnik, R., and Meissner, A. (2015b). Ground State Conditions Induce Rapid Reorganization of Core Pluripotency Factor Binding before Global Epigenetic Reprogramming. Cell stem cell.

Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011a). Open chromatin in pluripotency and reprogramming. Nature reviews Molecular cell biology *12*, 36-47.

Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011b). Open chromatin in pluripotency and reprogramming. Nature reviews Molecular cell biology *12*, 36-47.

Gaspar-Maia, A., Alajem, A., Polesso, F., Sridharan, R., Mason, M.J., Heidersbach, A., Ramalho-Santos, J., McManus, M.T., Plath, K., Meshorer, E.*, et al.* (2009). Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature *460*, 863-868.

Gasser, A.G.-S.a.S.M. (2016). On TADs and LADs- Spatial Control Over Gene Expression. Trends in Genetics *32*.

Gorkin, D.U., Leung, D., and Ren, B. (2014). The 3D genome in transcriptional regulation and pluripotency. Cell Stem Cell *14*, 762-775.

Grabole, N., Tischler, J., Hackett, J.A., Kim, S., Tang, F., Leitch, H.G., Magnusdottir, E., and Surani, M.A. (2013). Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. EMBO reports *14*, 629-637.

Graf, U., Casanova, E.A., Wyck, S., Dalcher, D., Gatti, M., Vollenweider, E., Okoniewski, M.J., Weber, F.A., Patel, S.S., Schmid, M.W.*, et al.* (2017). Pramel7 mediates ground-state pluripotency through proteasomal-epigenetic combined pathways. Nature cell biology *19*, 763-773.

Grob, A., Colleran, C., and McStay, B. (2011). UBF an Essential Player in Maintenance of Active NORs and Nucleolar Formation. Protein Rev *15*, 57-82.

Grozdanov, P., Georgiev, O., and Karagyozov, L. (2003). Complete sequence of the 45-kb mouse ribosomal DNA repeat: analysis of the intergenic spacer. Genomics *82*, 637-643.

Grummt, I., Maier, U., Ohrlein, A., Hassouna, N., and Bachellerie, J.P. (1985). Transcription of mouse rDNA terminates downstream of the 3' end of 28S RNA and involves interaction of factors with repeated sequences in the 3' spacer. Cell *43*, 801-810.

Gu, L., Frommel, S.C., Oakes, C.C., Simon, R., Grupp, K., Gerig, C.Y., Bar, D., Robinson, M.D., Baer, C., Weiss, M.*, et al.* (2015a). BAZ2A (TIP5) is involved in epigenetic alterations in prostate cancer and its overexpression predicts disease recurrence. Nature genetics *47*, 22-30.

Gu, L., Frommel, S.C., Oakes, C.C., Simon, R., Grupp, K., Gerig, C.Y., Bar, D., Robinson, M.D., Baer, C., Weiss, M.*, et al.* (2015b). BAZ2A (TIP5) is involved in epigenetic alterations in prostate cancer and its overexpression predicts disease recurrence. Nature genetics *47*, 22-30.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W.*, et al.* (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948-951.

Guetg, C., Lienemann, P., Sirri, V., Grummt, I., Hernandez-Verdun, D., Hottiger, M.O., Fussenegger, M., and Santoro, R. (2010). The NoRC complex mediates the heterochromatin formation and stability of silent rRNA genes and centromeric repeats. The EMBO journal *29*, 2135-2146.

Guetg, C., Scheifele, F., Rosenthal, F., Hottiger, M.O., and Santoro, R. (2012). Inheritance of silent rDNA chromatin is mediated by PARP1 via noncoding RNA. Molecular cell *45*, 790-800.

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y.*, et al.* (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. Cell *162*, 900-910.

Guzman-Ayala, M., Sachs, M., Koh, F.M., Onodera, C., Bulut-Karslioglu, A., Lin, C.J., Wong, P., Nitta, R., Song, J.S., and Ramalho-Santos, M. (2015). Chd1 is essential for the high transcriptional output and rapid growth of the mouse epiblast. Development *142*, 118-127.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yanez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B.*, et al.* (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell *169*, 693-707 e614.

Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C.*, et al.* (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell stem cell *13*, 360-369.

Hackett, J.A., and Surani, M.A. (2013). DNA methylation dynamics during the mammalian life cycle. Philosophical transactions of the Royal Society of London Series B, Biological sciences *368*, 20110328.

Hackett, J.A., and Surani, M.A. (2014). Regulatory principles of pluripotency: from the ground state up. Cell stem cell *15*, 416-430.

Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F.*, et al.* (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. Nature genetics *43*, 630-638.

Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S.*, et al.* (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell stem cell *6*, 479-491.

Hayashi, K., de Sousa Lopes, S.M.C., Tang, F., Lao, K., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell stem cell *3*, 391-401.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell *146*, 519-532.

He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L.*, et al.* (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science *333*, 1303-1307.

Hnisz, D., Day, D.S., and Young, R.A. (2016a). Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. Cell *167*, 1188-1200.

Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A.*, et al.* (2016b). Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science *351*, 1454-1458.

Ho, L., and Crabtree, G.R. (2010). Chromatin remodelling during development. Nature *463*, 474-484.

Ho, L., Miller, E.L., Ronan, J.L., Ho, W.Q., Jothi, R., and Crabtree, G.R. (2011). esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. Nature cell biology *13*, 903-913.

Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A.I., Ranish, J., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proceedings of the National Academy of Sciences of the United States of America *106*, 5181-5186.

Hoffman, M.M., Ernst, J., Wilder, S.P., Kundaje, A., Harris, R.S., Libbrecht, M., Giardine, B., Ellenbogen, P.M., Bilmes, J.A., Birney, E.*, et al.* (2013). Integrative annotation of chromatin elements from ENCODE data. Nucleic acids research *41*, 827-841.

Hota, S.K., and Bruneau, B.G. (2016). ATP-dependent chromatin remodeling during mammalian development. Development *143*, 2882-2897.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Huang, Y., and Rao, A. (2012). New functions for DNA modifications by TET-JBP. Nat Struct Mol Biol *19*, 1061-1064.

Hug, C.B., Grimaldi, A.G., Kruse, K., and Vaquerizas, J.M. (2017). Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. Cell *169*, 216-228.e219.

Ito, T., Levenstein, M.E., Fyodorov, D.V., Kutach, A.K., Kobayashi, R., and Kadonaga, J.T. (1999). ACF consists of two subunits, Acf1 and ISWI, that function cooperatively in the ATP-dependent catalysis of chromatin assembly. Genes & development *13*, 1529-1539.

Jadhav, U., Nalapareddy, K., Saxena, M., O'Neill, N.K., Pinello, L., Yuan, G.C., Orkin, S.H., and Shivdasani, R.A. (2016). Acquired Tissue-Specific Promoter Bivalency Is a Basis for PRC2 Necessity in Adult Cells. Cell *165*, 1389-1400.

Jonathan A. Beagan, M.T.D., Katelyn R. Titus, Linda Zhou, Zhendong Cao, Jingjing Ma, Caroline V. Lachanski, Daniel R. Gillis, Jennifer E. Phillips-Cremins (2017). YY1 and CTCF orchestrate a 3-D chromatin looping switch during early neural lineage commitment. Genome Research.

Jones, M.H., Hamana, N., Nezu, J., and Shimane, M. (2000). A novel family of bromodomain genes. Genomics *63*, 40-45.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews Genetics *13*, 484-492.

Joshi, O., Wang, S.Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P.J., Habibi, E., Shaik, J., Saeed, S., Handoko, L.*, et al.* (2015). Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. Cell stem cell *17*, 748-757.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Ke, Y., Xu, Y., Chen, X., Feng, S., Liu, Z., Sun, Y., Yao, X., Li, F., Zhu, W., Gao, L.*, et al.* (2017). 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. Cell *170*, 367-381 e320.

Kellum, R., and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. Cell *64*, 941-950.

Kidder, B.L., Palmer, S., and Knott, J.G. (2009). SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells. Stem cells (Dayton, Ohio) *27*, 317-328.

Kim, T.W., Kang, B.H., Jang, H., Kwak, S., Shin, J., Kim, H., Lee, S.E., Lee, S.M., Lee, J.H., Kim, J.H.*, et al.* (2015). Ctbp2 Modulates NuRD-Mediated Deacetylation of H3K27 and Facilitates PRC2-Mediated H3K27me3 in Active Embryonic Stem Cell Genes During Exit from Pluripotency. Stem cells (Dayton, Ohio) *33*, 2442-2455.

Kind, J., Pagie, L., de Vries, S.S., Nahidiazar, L., Dey, S.S., Bienko, M., Zhan, Y., Lajoie, B., de Graaf, C.A., Amendola, M.*, et al.* (2015). Genome-wide maps of nuclear lamina interactions in single human cells. Cell *163*, 134-147.

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S.S., Janssen, H., Amendola, M., Nolen, L.D., Bickmore, W.A., and van Steensel, B. (2013). Single-cell dynamics of genome-nuclear lamina interactions. Cell *153*, 178-192.

Kishikawa, S., Murata, T., Ugai, H., Yamazaki, T., and Yokoyama, K.K. (2003). Control elements of Dnmt1 gene are regulated in cell-cycle dependent manner. Nucleic acids research Supplement (2001), 307-308.

Kiyonari, H., Kaneko, M., Abe, S., and Aizawa, S. (2010). Three inhibitors of FGF receptor, ERK, and GSK3 establishes germline-competent embryonic stem cells of C57BL/6N mouse strain with high efficiency and stability. Genesis (New York, NY : 2000) *48*, 317-327.

Kojima, Y., Kaufman-Francis, K., Studdert, J.B., Steiner, K.A., Power, M.D., Loebel, D.A., Jones, V., Hor, A., de Alencastro, G., Logan, G.J.*, et al.* (2014). The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. Cell stem cell *14*, 107-120.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Kuhn, A., and Grummt, I. (1987). A novel promoter in the mouse rDNA spacer is active in vivo and in vitro. The EMBO journal *6*, 3487-3492.

Kumaran, R.I., and Spector, D.L. (2008). A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. J Cell Biol *180*, 51-65.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895-2902.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Langst, G., Becker, P.B., and Grummt, I. (1998). TTF-I determines the chromatin architecture of the active rDNA promoter. The EMBO journal *17*, 3135-3145.

Langst, G., Blank, T.A., Becker, P.B., and Grummt, I. (1997). RNA polymerase I transcription on nucleosomal templates: the transcription termination factor TTF-I induces chromatin remodeling and relieves transcriptional repression. The EMBO journal *16*, 760-768.

Lazzaro, M.A., and Picketts, D.J. (2001). Cloning and characterization of the murine Imitation Switch (ISWI) genes: differential expression patterns suggest distinct developmental roles for Snf2h and Snf2l. Journal of neurochemistry *77*, 1145-1156.

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., and Wutz, A. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. Genes & development *24*, 265-276.

Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A.*, et al.* (2013). Naive pluripotency is associated with global DNA hypomethylation. Nat Struct Mol Biol *20*, 311-316.

Leone, S., Bar, D., Slabber, C.F., Dalcher, D., and Santoro, R. (2017). The RNA helicase DHX9 establishes nucleolar heterochromatin, and this activity is required for embryonic stem cell differentiation. EMBO reports *18*, 1248-1262.

LeRoy, G., Loyola, A., Lane, W.S., and Reinberg, D. (2000). Purification and characterization of a human factor that assembles and remodels chromatin. The Journal of biological chemistry *275*, 14787-14790.

Lessard, J.A., and Crabtree, G.R. (2010). Chromatin regulatory mechanisms in pluripotency. Annu Rev Cell Dev Biol *26*, 503-532.

Lettice, L.A., Heaney, S.J., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Human molecular genetics *12*, 1725-1735.

Leung, A.K., Gerlich, D., Miller, G., Lyon, C., Lam, Y.W., Lleres, D., Daigle, N., Zomerdijk, J., Ellenberg, J., and Lamond, A.I. (2004). Quantitative kinetic analysis of nucleolar breakdown and reassembly during mitosis in live human cells. J Cell Biol *166*, 787-800.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li, E., and Zhang, Y. (2014). DNA methylation in mammals. Cold Spring Harbor perspectives in biology *6*, a019133.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589-595.

Li, J., Langst, G., and Grummt, I. (2006). NoRC-dependent nucleosome positioning silences rRNA genes. The EMBO journal *25*, 5735-5741.

Li, P., Tong, C., Mehrian-Shai, R., Jia, L., Wu, N., Yan, Y., Maxson, R.E., Schulze, E.N., Song, H., Hsieh, C.L.*, et al.* (2008). Germline competent embryonic stem cells derived from rat blastocysts. Cell *135*, 1299-1310.

Liang, G., and Zhang, Y. (2013). Embryonic stem cell and induced pluripotent stem cell: an epigenetic perspective. Cell research *23*, 49-69.

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic acids research *41*, e108.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J.*, et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nature genetics *38*, 431-440.

Lu, F., Liu, Y., Jiang, L., Yamaguchi, S., and Zhang, Y. (2014). Role of Tet proteins in enhancer activity and telomere elongation. Genes & development *28*, 2103-2119.

Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R.*, et al.* (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012-1025.

Malaguti, M., Nistor, P.A., Blin, G., Pegg, A., Zhou, X., and Lowell, S. (2013). Bone morphogenic protein signalling suppresses differentiation of pluripotent cells by maintaining expression of E-Cadherin. Elife *2*, e01197.

Manzo, M., Wirz, J., Ambrosi, C., Villasenor, R., Roschitzki, B., and Baubec, T. (2017). Isoform-specific localization of DNMT3A regulates DNA methylation fidelity at bivalent CpG islands. The EMBO journal *36*, 3421-3434.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A.*, et al.* (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell *149*, 590-604.

Martello, G., and Smith, A. (2014). The nature of embryonic stem cells. Annu Rev Cell Dev Biol *30*, 647-675.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Gottgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell stem cell *11*, 491-504.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proceedings of the National Academy of Sciences of the United States of America *78*, 7634-7638.

Marx, V. (2012). Epigenetics: Reading the second genomic code. Nature *491*, 143-147.

Mayer, C., Neubert, M., and Grummt, I. (2008). The structure of NoRC-associated RNA is crucial for targeting the chromatin remodelling complex NoRC to the nucleolus. EMBO reports *9*, 774-780.

Mayer, C., Schmitz, K.M., Li, J., Grummt, I., and Santoro, R. (2006). Intergenic transcripts regulate the epigenetic state of rRNA genes. Mol Cell *22*, 351-361.

McStay, B., and Grummt, I. (2008). The epigenetics of rRNA genes: from molecular to chromosome biology. Annu Rev Cell Dev Biol *24*, 131-157.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., *et al.* (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766-770.

Merkenschlager, M., and Nora, E.P. (2016). CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. Annu Rev Genomics Hum Genet *17*, 17-43.

Meshorer, E., and Misteli, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. Nature reviews Molecular cell biology *7*, 540-546.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Developmental cell *10*, 105-116.

Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J.B., Pagie, L., Kellis, M., Reinders, M., Wessels, L., and van Steensel, B. (2013). Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. Genome Res *23*, 270-280.

Miller, O.L., Jr., and Beatty, B.R. (1969). Visualization of nucleolar genes. Science *164*, 955-957.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.

Moss, T., Langlois, F., Gagnon-Kugler, T., and Stefanovsky, V. (2007). A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. Cellular and molecular life sciences : CMLS *64*, 29-49.

Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science *347*, 1017-1021.

Naughton, C., Avlonitis, N., Corless, S., Prendergast, J.G., Mati, I.K., Eijk, P.P., Cockroft, S.L., Bradley, M., Ylstra, B., and Gilbert, N. (2013). Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. Nat Struct Mol Biol *20*, 387-395.

Nemeth, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Peterfia, B., Solovei, I., Cremer, T., Dopazo, J., and Langst, G. (2010). Initial genomics of the human nucleolus. PLoS Genet *6*, e1000889.

Nemeth, A., and Grummt, I. (2018). Dynamic regulation of nucleolar architecture. Curr Opin Cell Biol *52*, 105-111.

Nemeth, A., and Langst, G. (2011). Genome organization in and around the nucleolus. Trends in genetics : TIG *27*, 149-156.

Nemeth, A., Strohner, R., Grummt, I., and Langst, G. (2004). The chromatin remodeling complex NoRC and TTF-I cooperate in the regulation of the mammalian rRNA genes in vivo. Nucleic acids research *32*, 4091-4099.

Nichols, J., Jones, K., Phillips, J.M., Newland, S.A., Roode, M., Mansfield, W., Smith, A., and Cooke, A. (2009). Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. Nature medicine *15*, 814-818.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell stem cell *4*, 487-492.

Nitiss, J.L. (2009). DNA topoisomerase II and its growing repertoire of biological functions. Nat Rev Cancer *9*, 327-337.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nature genetics *24*, 372-376.

Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944 e922.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Northcott, P.A., Lee, C., Zichner, T., Stutz, A.M., Erkek, S., Kawauchi, D., Shih, D.J., Hovestadt, V., Zapatka, M., Sturm, D.*, et al.* (2014). Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. Nature *511*, 428-434.

Nozaki, T., Imai, R., Tanbo, M., Nagashima, R., Tamura, S., Tani, T., Joti, Y., Tomita, M., Hibino, K., Kanemaki, M.T.*, et al.* (2017). Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. Molecular cell *67*, 282-293 e287.

O'Shaughnessy-Kirwan, A., Signolet, J., Costello, I., Gharbi, S., and Hendrich, B. (2015). Constraint of gene expression by the chromatin remodelling protein CHD4 facilitates lineage specification. Development *142*, 2586-2597.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247-257.

Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). Science *183*, 330-332.

Ong, C.T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. Nature reviews Genetics *15*, 234-246.

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. EMBO reports *10*, 1235-1241.

Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. Nature genetics *35*, 190-194.

Park, S.H., Park, S.H., Kook, M.C., Kim, E.Y., Park, S., and Lim, J.H. (2004). Ultrastructure of human embryonic stem cells and spontaneous and retinoic acid-induced differentiating cells. Ultrastructural pathology *28*, 229-238.

Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol Cell Biol *27*, 3769-3779.

Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. The EMBO journal *23*, 4061-4071.

Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W., Solovei, I., Brugman, W., Graf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M.*, et al.* (2010). Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. Molecular cell *38*, 603-613.

Petesch, S.J., and Lis, J.T. (2012). Overcoming the nucleosome barrier during transcript elongation. Trends in genetics : TIG *28*, 285-294.

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y.*, et al.* (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell *153*, 1281-1295.

Pikaard, C.S. (2000). The epigenetics of nucleolar dominance. Trends in genetics : TIG *16*, 495-500.

Pommier, Y., Leo, E., Zhang, H., and Marchand, C. (2010). DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. Chem Biol *17*, 421-433.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Rabl, C. (1885). Über Zellteilung. Morpholgisches Jahrbuch *10*, 214-330.

Racko, D., Benedetti, F., Dorier, J., and Stasiak, A. (2017). Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes. Nucleic acids research.

Rafiee, M.-R., Girardot, C., Sigismondo, G., and Krijgsveld, J. (2016). Expanding the Circuitry of Pluripotency by Selective Isolation of Chromatin-Associated Proteins. Molecular cell.

Ralf Strohner, R.S., Ingrid Grummt (2001). NoRC-a novel member of mammalian ISWI-containing chromatin remodeling machines. The EMBO journal.

Ramakrishnan, V. (1997). Histone structure and the organization of the nucleosome. Annual review of biophysics and biomolecular structure *26*, 83-112.

Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research *42*, W187-191.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nature protocols *8*, 2281-2308.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D.*, et al.* (2017). Cohesin Loss Eliminates All Loop Domains. Cell *171*, 305-320 e324.

Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature *452*, 243-247.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature *447*, 425-432.

Ricci, M.A., Manzo, C., Garcia-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. Cell *160*, 1145-1158.

Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L., and Wolffe, A.P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nature genetics *25*, 338-342.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nature Biotechnology.

Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology.

Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirier, F., Rigby, P.W., and Staudt, L.M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature *345*, 686-692.

Rothbart, S.B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A.I., Barsyte-Lovejoy, D., Martinez, J.Y., Bedford, M.T., Fuchs, S.M*., et al.* (2012). Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. Nat Struct Mol Biol *19*, 1155-1160.

Rothbart, S.B., and Strahl, B.D. (2014). Interpreting the language of histone and DNA modifications. Biochimica et biophysica acta *1839*, 627-643.

Sanborn, A.L., Rao, S.S., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J*., et al.* (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proceedings of the National Academy of Sciences of the United States of America *112*, E6456-6465.

Santoro, R. (2005). The silence of the ribosomal RNA genes. Cellular and molecular life sciences : CMLS *62*, 2067-2079.

Santoro, R. (2011). The Epigenetics of the Nucleolus: Structure and Function of Active and Silent Ribosomal RNA Genes. Protein Rev *15*, 57-82.

Santoro, R., and Grummt, I. (2001). Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription. Molecular cell *8*, 719-725.

Santoro, R., and Grummt, I. (2005). Epigenetic mechanism of rRNA gene silencing: temporal order of NoRC-mediated histone modification, chromatin remodeling, and DNA methylation. Mol Cell Biol *25*, 2539-2546.

Santoro, R., Li, J., and Grummt, I. (2002). The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. Nature genetics *32*, 393-396.

Santoro, R., Schmitz, K.M., Sandoval, J., and Grummt, I. (2010). Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans. EMBO reports *11*, 52-58.

Sarma, K., and Reinberg, D. (2005). Histone variants meet their match. Nature reviews Molecular cell biology *6*, 139-149.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nature medicine *10*, 55-63.

Savic, N., Bar, D., Leone, S., Frommel, S.C., Weber, F.A., Vollenweider, E., Ferrari, E., Ziegler, U., Kaech, A., Shakhova, O*., et al.* (2014). lncRNA Maturation to Initiate Heterochromatin Formation in the Nucleolus Is Required for Exit from Pluripotency in ESCs. Cell stem cell *15*, 720-734.

Savić, N., Bär, D., Leone, S., Frommel, S.C., Weber, F.A., Vollenweider, E., Ferrari, E., Ziegler, U., Kaech, A., Shakhova, O*., et al.* (2014). lncRNA Maturation to Initiate Heterochromatin Formation in the Nucleolus Is Required for Exit from Pluripotency in ESCs. In Cell Stem Cell, pp. 720-734.

Scheer, U., and Hock, R. (1999). Structure and function of the nucleolus. Curr Opin Cell Biol *11*, 385-390.

Schlesinger, S., Selig, S., Bergman, Y., and Cedar, H. (2009). Allelic inactivation of rDNA loci. Genes & development *23*, 2437-2447.

Schmid, M.W., and Grossniklaus, U. (2015). Rcount: simple and flexible RNA-Seq read counting. Bioinformatics *31*, 436-437.

Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M*., et al.* (2015a). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. Nature genetics.

Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M*., et al.* (2015b). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. Nat Genet *47*, 1179-1186.

Scholer, H.R., Dressler, G.R., Balling, R., Rohdewohld, H., and Gruss, P. (1990). Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. The EMBO journal *9*, 2185-2195.

Schwarz, B.A., Bar-Nur, O., Silva, J.C., and Hochedlinger, K. (2014). Nanog is dispensable for the generation of induced pluripotent stem cells. Current biology : CB *24*, 347-350.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., C, H.H., Mirny, L.*, et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51-56.

Shah, M.Y., and Licht, J.D. (2011). DNMT3A mutations in acute myeloid leukemia. Nature genetics *43*, 289-290.

Sharif, J., Endo, T.A., Nakayama, M., Karimi, M.M., Shimada, M., Katsuyama, K., Goyal, P., Brind'Amour, J., Sun, M.A., Sun, Z.*, et al.* (2016). Activation of Endogenous Retroviruses in Dnmt1(-/-) ESCs Involves Disruption of SETDB1-Mediated Repression by NP95 Binding to Hemimethylated DNA. Cell stem cell *19*, 81-94.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K.*, et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908-912.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annual review of biochemistry *81*, 65-95.

Shipony, Z., Mukamel, Z., Cohen, N.M., Landan, G., Chomsky, E., Zeliger, S.R., Fried, Y.C., Ainbinder, E., Friedman, N., and Tanay, A. (2014). Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells. Nature *513*, 115-119.

Silva, J., and Smith, A. (2008). Capturing pluripotency. Cell *132*, 532-536.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688-690.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nature reviews Genetics *14*, 204-220.

Song, J., Rechkoblit, O., Bestor, T.H., and Patel, D.J. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. Science *331*, 1036-1040.

Stadhouders, R., Vidal, E., Serra, F., Di Stefano, B., Le Dily, F., Quilez, J., Gomez, A., Collombet, S., Berenguer, C., Cuartero, Y.*, et al.* (2018). Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. Nature genetics.

Stopka, T., and Skoultchi, A.I. (2003). The ISWI ATPase Snf2h is required for early mouse development. Proceedings of the National Academy of Sciences of the United States of America *100*, 14097-14102.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Strohner, R., Nemeth, A., Jansa, P., Hofmann-Rohrer, U., Santoro, R., Langst, G., and Grummt, I. (2001). NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines. The EMBO journal *20*, 4892-4900.

Subramanian, V., Mazumder, A., Surface, L.E., Butty, V.L., Fields, P.A., Alwan, A., Torrey, L., Thai, K.K., Levine, S.S., Bathe, M.*, et al.* (2013). H2A.Z acidic patch couples chromatin dynamics to regulation of gene expression programs during ESC differentiation. PLoS Genet *9*, e1003725.

Suzuki, S., Nozawa, Y., Tsukamoto, S., Kaneko, T., Manabe, I., Imai, H., and Minami, N. (2015). CHD1 acts via the Hmgpi pathway to regulate mouse early embryogenesis. Development *142*, 2375-2384.

Taberlay, P.C., Kelly, T.K., Liu, C.C., You, J.S., De Carvalho, D.D., Miranda, T.B., Zhou, X.J., Liang, G., and Jones, P.A. (2011). Polycomb-repressed genes have permissive enhancers that initiate reprogramming. Cell *147*, 1283-1294.

Tan-Wong, S.M., Zaugg, J.B., Camblong, J., Xu, Z., Zhang, D.W., Mischo, H.E., Ansari, A.Z., Luscombe, N.M., Steinmetz, L.M., and Proudfoot, N.J. (2012). Gene loops enhance transcriptional directionality. Science *338*, 671-675.

Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B.*, et al.* (2015). CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell *163*, 1611-1627.

Tate, P.H., and Bird, A.P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. Curr Opin Genet Dev *3*, 226-231.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat Struct Mol Biol *14*, 1025-1040.

Tee, W.W., and Reinberg, D. (2014). Chromatin features and the epigenetic regulation of pluripotency states in ESCs. Development *141*, 2376-2390.

ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nature cell biology *13*, 1070-1075.

Thakurela, S., Garding, A., Jung, J., Schubeler, D., Burger, L., and Tiwari, V.K. (2013). Gene regulation and priming by topoisomerase IIalpha in embryonic stem cells. Nature communications *4*, 2478.

Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L., and Melton, D.W. (1989). Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. Cell *56*, 313-321.

Tiwari, V.K., Burger, L., Nikoletopoulou, V., Deogracias, R., Thakurela, S., Wirbelauer, C., Kaut, J., Terranova, R., Hoerner, L., Mielke, C.*, et al.* (2012). Target genes of Topoisomerase IIbeta regulate neuronal survival and are defined by their chromatin state. Proc Natl Acad Sci U S A *109*, E934-943.

Torres-Padilla, M.E., and Zernicka-Goetz, M. (2006). Role of TIF1alpha as a modulator of embryonic transcription in the mouse zygote. J Cell Biol *174*, 329-338.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. Development *135*, 909-918.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R.*, et al.* (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes to cells : devoted to molecular & cellular mechanisms *11*, 805-814.

Uuskula-Reimand, L., Hou, H., Samavarchi-Tehrani, P., Rudan, M.V., Liang, M., Medina-Rivera, A., Mohammed, H., Schmidt, D., Schwalie, P., Young, E.J.*, et al.* (2016). Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. Genome Biol *17*, 182.

van Koningsbruggen, S., Gierlinski, M., Schofield, P., Martin, D., Barton, G.J., Ariyurek, Y., den Dunnen, J.T., and Lamond, A.I. (2010). High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. Molecular biology of the cell *21*, 3735-3748.

van Steensel, B., and Belmont, A.S. (2017). Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell *169*, 780-791.

Voigt, P., Tee, W.W., and Reinberg, D. (2013). A double take on bivalent promoters. Genes & development *27*, 1318-1338.

von Meyenn, F., Iurlaro, M., Habibi, E., Liu, N.Q., Salehzadeh-Yazdi, A., Santos, F., Petrini, E., Milagre, I., Yu, M., Xie, Z.*, et al.* (2016). Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. Molecular cell.

Wang, L., Du, Y., Ward, J.M., Shimbo, T., Lackford, B., Zheng, X., Miao, Y.L., Zhou, B., Han, L., Fargo, D.C.*, et al.* (2014). INO80 facilitates pluripotency gene activation in embryonic stem cell self-renewal, reprogramming, and blastocyst development. Cell stem cell *14*, 575-591.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. Nature reviews Molecular cell biology.

Weiner, A., Lara-Astiaso, D., Krupalnik, V., Gafni, O., David, E., Winter, D.R., Hanna, J.H., and Amit, I. (2016). Co-ChIP enables genome-wide mapping of histone mark co-occurrence at single-molecule resolution. Nat Biotechnol *34*, 953-961.

Wiechens, N., Singh, V., Gkikopoulos, T., Schofield, P., Rocha, S., and Owen-Hughes, T. (2016). The Chromatin Remodelling Enzymes SNF2H and SNF2L Position Nucleosomes adjacent to CTCF and Other Transcription Factors. PLoS Genet *12*, e1005940.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343-348.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature *336*, 684-687.

Wingett, S., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., and Andrews, S. (2015). HiCUP: pipeline for mapping and processing Hi-C data. F1000Research *4*, 1310.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nature cell biology *13*, 838-845.

Wu, H., and Zhang, Y. (2014). Reversing DNA methylation: mechanisms, genomics, and biological functions. Cell *156*, 45-68.

Yamada, T., Yang, Y., Hemberg, M., Yoshida, T., Cho, H.Y., Murphy, J.P., Fioravante, D., Regehr, W.G., Gygi, S.P., Georgopoulos, K.*, et al.* (2014). Promoter decommissioning by the NuRD chromatin remodeling complex triggers synaptic connectivity in the mammalian brain. Neuron *83*, 122-134.

Yang, F., Deng, X., Ma, W., Berletch, J.B., Rabaia, N., Wei, G., Moore, J.M., Filippova, G.N., Xu, J., Liu, Y.*, et al.* (2015). The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. Genome Biol *16*, 52.

Yeo, J.C., Jiang, J., Tan, Z.Y., Yim, G.R., Ng, J.H., Goke, J., Kraus, P., Liang, H., Gonzales, K.A., Chong, H.C.*, et al.* (2014). Klf2 is an essential factor that sustains ground state pluripotency. Cell stem cell *14*, 864-872.

Yi, F., Pereira, L., Hoffman, J.A., Shy, B.R., Yuen, C.M., Liu, D.R., and Merrill, B.J. (2011). Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. Nature cell biology *13*, 762-770.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281-292.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-523.

Yip, D.J., Corcoran, C.P., Alvarez-Saavedra, M., DeMaria, A., Rennick, S., Mears, A.J., Rudnicki, M.A., Messier, C., and Picketts, D.J. (2012). Snf2l regulates Foxg1-dependent progenitor cell expansion in the developing brain. Developmental cell *22*, 871-878.

Yokochi, T., and Robertson, K.D. (2002). Preferential methylation of unmethylated DNA by Mammalian de novo DNA methyltransferase Dnmt3a. The Journal of biological chemistry *277*, 11735-11745.

Young, R.A. (2011). Control of the embryonic stem cell state. Cell *144*, 940-954.

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics *25*, 1952-1958.

Zentner, G.E., Balow, S.A., and Scacheri, P.C. (2014). Genomic characterization of the mouse ribosomal DNA locus. G3 (Bethesda) *4*, 243-254.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhao, H., Sun, Z., Wang, J., Huang, H., Kocher, J.P., and Wang, L. (2014). CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics *30*, 1006-1007.

Zhao, X.D., Han, X., Chew, J.L., Liu, J., Chiu, K.P., Choo, A., Orlov, Y.L., Sung, W.K., Shahab, A., Kuznetsov, V.A.*, et al.* (2007). Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell stem cell *1*, 286-298.

Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nature reviews Genetics *12*, 7-18.

Zhou, Y., and Grummt, I. (2005). The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing. Current biology : CB *15*, 1434-1438.

Zhou, Y., Santoro, R., and Grummt, I. (2002). The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. The EMBO journal *21*, 4632-4640.

Zhou, Y., Schmitz, K.M., Mayer, C., Yuan, X., Akhtar, A., and Grummt, I. (2009). Reversible acetylation of the chromatin remodelling complex NoRC is required for non-coding RNA-dependent silencing. Nature cell biology *11*, 1010-1016.

Zuin, J., Franke, V., van Ijcken, W.F., van der Sloot, A., Krantz, I.D., van der Reijden, M.I., Nakato, R., Lenhard, B., and Wendt, K.S. (2014). A cohesin-independent role for NIPBL at promoters provides insights in CdLS. PLoS Genet *10*, e1004153.

6. Curriculum Vitae

Personal Data

Name	Dalcher Damian
Address	Kurvenstrasse 9
	CH-8600 Dübendorf
Telephone	+41 79 307 38 29
Email	damian.dalcher@gmail.com
Date of birth	22 November 1987
Home town	Trubschachen (BE)
Nationality	Swiss

Education

Since 5/2014	**Molecular Life Science PhD program, Zurich** University of Zurich, Switzerland *"Chromatin and Epigenetic Regulation in Ground State Pluripotency"*
8/2013-5/2014	**Internship as scientific associate in Molecular Cancer Research** University of Zurich, Switzerland *"Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells"*
9/2011-6/2013	**Master of Science in Molecular Life Science (mark 6.0)** University of Bern, Switzerland Master thesis: "*Histone mRNA 3' end formation: Insight into two regulatory mechanisms*" (mark 6.0)
8/2008-6/2011	**Bachelor of Science in Cell Biology (mark 5.5)** University of Bern, Switzerland Bachelor thesis: "*U7 snRNP affinity purification using a MS2-MBP method"* (mark 6.0)
8/2003-6/2007	**Matura (mark 5.1)** Seelandgymnasium Biel, Switzerland Major: Biology and Chemistry

Awards

2015	**Poster prize at for the overall best poster at the Molecular Life Sciences PhD student retreat**
2015	**Forschungskredit: Candoc, University of Zurich** Funding for a yearly PhD salary

Publications

Graf, U., Casanova, E.A., Wyck, S., **Dalcher, D.**, Gatti, M., Vollenweider, E., Okoniewski, M.J., Weber, F.A., Patel, S.S., Schmid, M.W*., et al.* (2017). Pramel7 mediates ground-state pluripotency through proteasomal-epigenetic combined pathways. **Nature cell biology** *19*, 763-773.

Leone, S., Bar, D., Slabber, C.F., **Dalcher, D.,** and Santoro, R. (2017). The RNA helicase DHX9 establishes nucleolar heterochromatin, and this activity is required for embryonic stem cell differentiation. **EMBO reports** *18*, 1248-1262.

Kresoja-Rakic, J., Kapaklikaya, E., Ziltener, G., **Dalcher, D**., Santoro, R., Christensen, B.C., Johnson, K.C., Schwaller, B., Weder, W., Stahel, R.A*., et al.* (2016). Identification of cis- and trans-acting elements regulating calretinin expression in mesothelioma cells. **Oncotarget** *7*, 21272-21286.

Zellweger, R.*, **Dalcher, D.***, Mutreja, K., Berti, M., Schmid, J.A., Herrador, R., Vindigni, A., and Lopes, M. (2015). Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. **J Cell Biol** *208*, 563-579.
** These authors contributed equally to this work*

7. Acknowledgments

Throughout my PhD I have encountered numerous challenges and obstacles, which I could not have overcome without the generous help of many people.

First of all, I would like to express my outmost gratitude to my PhD supervisor PD Dr. Raffaella Santoro for giving me the honor of being a part of her research group. Her unrestricted passion for science is absolutely remarkable and encouraged me various times to keep going with my project at tough moments. Doubtlessly, my PhD project would not have reached the same level and quality without her highly competent scientific guidance.

I am also very grateful to my thesis committee members Prof. Dr. Dr. Michael O. Hottiger, Prof. Dr. Constance Ciaudo, PD Dr. Paolo Cinelli and Dr. Marcus Buschbeck for their support and constructive discussions during my yearly PhD thesis committee meetings. Special thanks go to Prof. Dr. Constance Ciaudo and PD Dr. Paolo Cinelli for the generous sharing of various embryonic stem cell lines and for the helpful advice in culturing these delicate cells.

Four years of PhD work would not have been possible without the wonderful atmosphere in our lab and institute. Thus, I would like to thank all the current and former “Santoros”, namely my coffee buddy Dr. Sergio Leone, the qPCR pipetting machine Dominik Bär, the cow creator Sarah Wyck, our multinational Mexican/Puerto Rican/Costa Rican Rodrigo Pena-Hernandez, the fragile lady Dr. Isabella Zanini, my two Polish friends Karolina Pietrzak and Marcin Damian Roganowicz, our Italian connection including Cristiana Bersaglieri, Giorgia Rizzari and Francesco Ruberto, my Dobby Julia Kuhn, our lively friend from Sri Lanka Meneka Ruvi Rupasinghe, the two eager-to-learn Erik Slabber and Juliana Bizzarro and the three pioneers Dr. Eva Vollenweider, Dr. Sandra Frommel and Dr. Natasa Savic for the countless scientific and non-scientific discussion and for making this time a great experience and a lot of fun. At this point I would also like to point out that the scientific spirit was not only exceptional within the Santoro group but also among the entire institute. Accordingly, I am very thankful to each DMMD and CABMM member for all the work-related discussions but also for all the coffee breaks and after-work beers. I would definitely not have survived this time without being a part of this great team.

Furthermore, I have largely benefited from the extraordinary scientific surrounding given by the University of Zurich and the Life Science Graduate School Zurich. Special thanks also go to Jelena Kühn-Georgijevic from the Functional Genomics Center Zurich for the all the work on next generation sequencing. I'm also highly grateful for being awarded with the Forschungskredit Candoc, which greatly helped financially supporting this work.

I want to also deeply thank my family and friends. My guys back home never let a doubt about fièr d'être Biennois! Further, I cannot thank Céline enough for having my back at any time and for the impossible achievement of making Bern feel like home. Last but definitely not least my deepest gratitude goes to my parents Marianne and Allan for the priceless support and advice over all these years. THANK YOU!